```
OM of: US-09-613-508-4 to: N_Geneseq_032802:*   out_format : pfs

Date: Aug 15, 2002 10:21 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=x1h
 -Q=/cgn2_1/USPTO_spool/US09613508/runat_14082002_112714_7815/app_query.fasta_1.701
 -DB=N_Geneseq_032802 -QFMT=fastap -SUFFIX=olig.rng -GAPOP=4.500
 -GAPEXT=0.050 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000 -YGAPEXT=60.000
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=oligo
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality
 -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext
 -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US09613508_@CGN1_1_190 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-613-508-4
Query length: 277
Database: N_Geneseq_032802:*
Database sequences: 1736436
Database length: 858457221
Search time (sec): 265.980000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000


score_list:
Sequence                Strd Orig    ZScore    EScore Len  ! Documentation  ..
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13075 +  277.00 5024.11  1.7e-2
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT33567 +  277.00 5021.73  2.3e-27
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV05471 +  277.00 5021.73  2.3e-27
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT15277 +  277.00 5021.73  2.3e-27
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAC64043 +  277.00 5016.05  4.8e-27
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13078 +  275.00 4972.89  1.2e-2
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13079 +  275.00 4972.60  1.3e-2
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13080 +  275.00 4972.34  1.3e-2
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13081 +  275.00 4972.09  1.3e-2
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81219 +  189.00 3421.68  3.1e-18
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAS10503 +  189.00 3413.36  8.9e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03910 +  189.00 3413.36  8.9e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13084 +  180.00 3242.15  3.1e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV32608 +  176.00 3183.65  5.5e-16
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV32615 +  176.00 3183.65  5.5e-16
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81214 +  175.00 3166.43  5.0e-16
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81215 +  147.00 2656.69  1.2e-13
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03918 +  146.00 2639.12  1.2e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAH26878 +  146.00 2628.07  4.9e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAC01186 +   67.00 1204.73  9.3e-59
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAI66510 +   36.00  635.66  4.7e-2
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAS10510 +   35.00  614.25  7.2e-2
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV13959 +   30.00  523.91  7.8e-21
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV13960 +   30.00  523.91  7.8e-21
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAL27190 -   16.00  291.71  6.7e-0
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAF90397 +   15.00  260.72  3.6e-0
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13076 +   14.00  255.43  7.0e-0
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13082 +   14.00  255.01  7.4e-0
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13077 -   13.00  237.52  7.0e-0
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAV33997 +   11.00  200.38  0.0082
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAV33993 +   11.00  200.38  0.0082
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX78269 +   11.00  196.96  0.0127
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX78257 +   11.00  196.96  0.0127
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAC63843 +   11.00  196.96  0.0127
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAC63855 +   11.00  196.96  0.0127
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAA46609 +   11.00  196.96  0.0127
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAZ25254 +   10.00  183.86  0.0681
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX78265 +   10.00  178.34  0.1382
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX78253 +   10.00  178.34  0.13
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAC63839 +   10.00  178.34  0.13
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAC63851 +   10.00  178.34  0.13
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAA46593 +   10.00  178.34  0.13
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAA46605 +   10.00  178.34  0.13
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81223 +   10.00  161.56  1.19
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL26207 +   10.00  161.45  1.2

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13075

seq_documentation_block:
ID   AAD13075 standard; cDNA; 834 BP.
XX
AC   AAD13075;
XX
DT   16-OCT-2001  (first entry)
XX
DE   Mammalia caspase-3 reporter moiety cDNA.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance;
KW   caspase-3; ss.
XX
OS   Mammalia.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain  -
XX
PS   Example 4; Page 131; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is mammalia caspase-3 cDNA which is used as a reporter moiety in
CC   the exemplification of the invention.
XX
SQ   Sequence 834 BP; 274 A; 147 C; 179 G; 234 T; 0 other;


alignment_scores:
            Quality:  277.00              Length:    277
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
```

```
US-09-613-508-4 x AAD13075   ..

Align seg 1/1  to: AAD13075  from: 1  to: 834

      1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      1 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 50

     17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 100

     34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    101 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 150

     51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    151 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 200

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    201 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 250

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    251 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 300

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    301 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 350

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    351 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 400

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    401 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 450

    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    451 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 500

    167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    501 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 550

    184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    551 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 600

    201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    601 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 650

    217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    651 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 700

    234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    701 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 750

    251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    751 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 800

    267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
        ||||||||||||||||||||||||||||||
    801 CATGCTCACAAAAGAACTCTATTTTTATCAC 831

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT33567
```

```
seq_documentation_block:
ID   AAT33567 standard; cDNA; 1158 BP.
XX
AC   AAT33567;
XX
DT   06-DEC-1996  (first entry)
XX
DE   Pro-Yama cDNA.
XX
KW   Yama; CrmA; apoptosis; Fas receptor; T-lymphocyte; AIDS; HIV;
KW   human immunodeficiency virus; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             224..1057
FT                   /*tag=  a
XX
PN   WO9625501-A2.
XX
PD   22-AUG-1996.
XX
PF   09-FEB-1996;   96WO-US01882.
XX
PR   01-JUN-1995;   95US-0457731.
PR   13-FEB-1995;   95US-0389812.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Dixit VM;
XX
DR   WPI; 1996-393402/39.
DR   P-PSDB; AAW00677.
XX
PT   New nucleic acid encoding Yama protein or CrmA mutant - useful for
PT   modulating apoptosis, maintaining T cell viability in AIDS patients
PT   and for drug screening
XX
PS   Disclosure; Fig 1; 118pp; English.
XX
CC   A cDNA clone (AAT33567) codes for pro-Yama (AAW00677), a zymogen which
CC   upon activation cleaves PARP to an 85 kDa form.  Activated Yama has
CC   the ability to modulate cellular function associated with the Fas
CC   receptor pathway such as Fas-associated apoptosis.  This activity is
CC   inhibited by CrmA (see also AAW00678).  The cDNA was isolated from a
CC   cDNA library prepd. from human umbilical vein endothelial cells
CC   treated with tumour necrosis factor and cycloheximide.  It can be
CC   used to produce Yama in a host cell for use in drug screening.
CC   Yama and anti-Yama protein nucleic acids, or antisense RNA, can be
CC   used to modulate cellular functions such as apoptosis and immune
CC   disorders mediated by the Fas pathway.
XX
SQ   Sequence 1158 BP; 328 A; 220 C; 275 G; 335 T; 0 other;


alignment_scores:
            Quality:  277.00              Length:    277
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AAT33567   ..

 Align seg 1/1  to: AAT33567  from: 1  to: 1158

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     224 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 273

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     274 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 323
```

```
      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     324 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 373

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     374 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 423

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     424 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 473

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     474 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 523

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     524 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 573

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     574 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 623

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     624 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 673

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     674 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 723

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     724 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 773

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     774 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 823

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     824 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 873

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     874 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 923

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     924 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 973

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     974 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1023

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         |||||||||||||||||||||||||||||
    1024 CATGCTCACAAAAGAACTCTATTTTTATCAC 1054
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV05471

seq_documentation_block:

```
ID   AAV05471 standard; DNA; 1158 BP.
XX
AC   AAV05471;
XX
DT   02-JUL-1998  (first entry)
XX
DE   Nucleic acid encoding a protein designated YAMA.
XX
KW   CrmA; tumour necrosis factor receptor; TNF-R; U1-70; inhibition;
KW   apoptosis; activation; U1-70 path; apoptosis path; identification;
KW   modulation; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             224..1057
FT                   /*tag= a
XX
PN   JP09299077-A.
XX
PD   25-NOV-1997.
XX
PF   27-JAN-1997;   97JP-0012932.
XX
PR   09-FEB-1996;   96US-0591605.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
DR   WPI; 1998-056551/06.
DR   P-PSDB; AAW41688.
XX
PT   Controlling apoptosis by inhibiting activation of the U1-70 pathway
PT   - by introduction of nucleic acid encoding protein with CrmA
PT   biological activity into the cell
XX
PS   Disclosure; Fig 1; 48pp; Japanese.
XX
CC   The present sequence encodes a non-naturally occurring protein designated
CC   YAMA. The specification describes a novel method of controlling cell
CC   functions controlled by the tumour necrosis factor receptor (TNF-R)
CC   pathway containing U1-70. The method comprises introducing a nucleic
CC   acid molecule encoding CrmA into a cell, and culturing the cell under
CC   suitable conditions for transcription and translation of the CrmA nucleic
CC   acid. Method for preventing and inhibiting apoptosis by inhibiting
CC   activation of the U1-70 path, identifying the chemical related to the
CC   apoptosis path, and screening a candidate chemical having biological
CC   function in the apoptosis path are also described. The methods can be
CC   used to identify substances which modulate apoptosis in cells controlled
CC   by the TNF-R pathway.
XX
SQ   Sequence 1158 BP; 328 A; 220 C; 275 G; 335 T; 0 other;
```

alignment_scores:

```
            Quality:  277.00              Length:    277
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000
```

alignment_block:

US-09-613-508-4 x AAV05471 ..

Align seg 1/1 to: AAV05471 from: 1 to: 1158

```
       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     224 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 273

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     274 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 323

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     324 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 373

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     374 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 423

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     424 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 473
```

```
     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    474 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 523

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    524 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 573

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    574 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 623

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    624 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 673

    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    674 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 723

    167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    724 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 773

    184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    774 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 823

    201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    824 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 873

    217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    874 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 923

    234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    924 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 973

    251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    974 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1023

    267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
        |||||||||||||||||||||||||||||||
   1024 CATGCTCACAAAAGAACTCTATTTTTATCAC 1054
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT15277

seq_documentation_block:

```
ID   AAT15277 standard; cDNA; 1159 BP.
XX
AC   AAT15277;
XX
DT   28-OCT-1996  (first entry)
XX
DE   Human interleukin-I-converting enzyme-like apoptosis protease-4 cDNA.
XX
KW   ICE-LAP-4; interleukin-I-converting enzyme-like apoptosis protease;
KW   enzyme; Alzheimer's disease; Parkinson's disease; septic shock;
KW   head injury; rheumatoid arthritis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             115..1137
FT                   /*tag= a
XX
PN   WO9613603-A1.
XX
PD   09-MAY-1996.
XX
PF   06-JUN-1995;   95WO-US07235.
XX
PR   01-NOV-1994;   94US-0334251.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Hastings GA,  He W,  Hudson PL,  Rosen CA;
XX
DR   WPI; 1996-239509/24.
DR   P-PSDB; AAR95831.
XX
PT   Human ICE-LAP-3 and -4 DNA and protein  - useful in the diagnosis
PT   and treatment of Alzheimer's disease, Parkinson's disease,
PT   rheumatoid arthritis, septic shock and head injury
XX
PS   Claim 1; Page 44; 67pp; English.
XX
CC   This sequence was derived from a cDNA library from human endometrial
CC   tumuor and may be used for the recombinant production of the ICE-LAP-
CC   4 protein e.g. in COS cells.  The encoded protein may be used
CC   therapeutically, e.g. as an antitumor or antiviral agent and to
CC   control embryonic development and tissue homeostasis.  The protein
CC   can also be used to treat immunosuppression disorders, such as AIDS,
CC   by targeting virus infected cells for cell death.  The DNA may find
CC   use in gene therapy applications.
XX
SQ   Sequence 1159 BP; 329 A; 220 C; 276 G; 334 T; 0 other;
```

alignment_scores:

```
            Quality:  277.00              Length:    277
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000
```

alignment_block:

US-09-613-508-4 x AAT15277 ..

Align seg 1/1 to: AAT15277 from: 1 to: 1159

```
      1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    223 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 272

     17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    273 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 322

     34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    323 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 372

     51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    373 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 422

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    423 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 472

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    473 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 522

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    523 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 572

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    573 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 622

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
```

```
      623 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 672

      151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
      673 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 722

      167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
      723 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 772

      184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
      773 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 822

      201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
      823 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 872

      217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
      873 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 922

      234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
      923 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 972

      251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
      973 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1022

      267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
     1023 CATGCTCACAAAAGAACTCTATTTTTATCAC 1053
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAC64043

seq_documentation_block:

```
ID   AAC64043 standard; DNA; 2538 BP.
XX
AC   AAC64043;
XX
DT   14-FEB-2001  (first entry)
XX
DE   pYNGCas3p10p35 insert encoding human caspase-3 and AcNPV p35 protein.
XX
KW   Human caspase-3; AcNPV p35 protein; BmNPV p10 promoter;
KW   polyhedron promoter; baculovirus vector; apoptosis inhibitor protein;
KW   recominant caspase production; insect cell; ds.
XX
OS   Chimeric - Homo sapiens.
OS   Chimeric - Autographa californica nuclear polyhedrosis virus.
OS   Chimeric - Bombyx mori nuclear polyhedrosis virus.
XX
PN   JP2000253882-A.
XX
PD   19-SEP-2000.
XX
PF   10-MAR-1999;   99JP-0062732.
XX
PR   10-MAR-1999;   99JP-0062732.
XX
PA   (KATA ) KATAKURA IND CO LTD.
XX
DR   WPI; 2000-658811/64.
XX
PT   Baculovirus vector used for the preparation of caspase protein,
PT   comprises a transfer vector plasmid containing a caspase
PT   protein-producing gene -
XX
PS   Example 1; Page 7-8; 10pp; Japanese.
XX
CC   The invention relates to a novel baculovirus vector used for the
CC   preparation of a caspase protein, and containing a recombinant transfer
CC   vector plasmid comprising a gene encoding a caspase and a gene encoding
CC   an apoptosis-inhibiting protein. The invention also encompasses the
CC   preparation of caspase protein, comprising transfecting the novel
CC   baculovirus vector into an insect or an insect cell; breeding the
CC   insect or culturing the cell; and collecting the caspase protein from
CC   the insect or insect cell. The present sequence represents the expression
CC   cassette insert from a transfer vector (pYNGCas3p10p35) produced in an
CC   exemplification of the invention. This insert comprises a mature human
CC   caspase-3 gene under the control of a polyhedron promoter, and an
CC   Autographa californica nuclear polyhedrosis virus (AcNPV) p35 gene under
CC   the control of a Bombyx mori nuclear polyhedrosis virus (BmNPV) p10
CC   promoter.
XX
SQ   Sequence 2538 BP; 865 A; 446 C; 506 G; 721 T; 0 other;
```

alignment_scores:

```
            Quality:  277.00              Length:    277
              Ratio:  1.000                 Gaps:      0
Percent Similarity: 100.000   Percent Identity: 100.000
```

alignment_block:

US-09-613-508-4 x AAC64043 ..

Align seg 1/1 to: AAC64043 from: 1 to: 2538

```
        1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
      136 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 185

       17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
      186 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 235

       34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
      236 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 285

       51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
      286 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 335

       67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
      336 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 385

       84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
      386 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 435

      101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
      436 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 485

      117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
      486 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 535

      134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
      536 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 585

      151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
      586 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 635

      167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
      636 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 685

      184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
```

```
     686 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 735

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     736 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 785

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     786 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 835

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     836 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 885

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     886 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 935

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     936 CATGCTCACAAAAGAACTCTATTTTTATCAC 966

seq_name: /SIDSl/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13078

seq_documentation_block:
ID   AAD13078 standard; DNA; 6436 BP.
XX
AC   AAD13078;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-1XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-1XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain -
XX
PS   Example 4; Page 133-134; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-1XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domain-Mammalia caspase-3 fusion protein.
XX
SQ   Sequence 6436 BP; 1577 A; 1606 C; 1607 G; 1646 T; 0 other;


alignment_scores:
           Quality:  275.00              Length:    275
             Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AAD13078   ..

 Align seg 1/1  to: AAD13078  from: 1  to: 6436

       3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1144 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAA 1193

      19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1194 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1243

      36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1244 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 1293

      53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1294 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 1343

      69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1344 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 1393

      86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1394 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1443

     103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1444 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 1493

     119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1494 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 1543

     136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1544 TGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 1593

     153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1594 GGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 1643

     169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1644 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 1693

     186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1694 AAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 1743

     203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1744 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 1793

     219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 1794 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 1843

  236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1844 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 1893

  253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1894 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 1943

  269 uThrLysGluLeuTyrPheTyrHis 277
      |||||||||||||||||||||||||
 1944 CACAAAAGAACTCTATTTTTATCAC 1968

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13079

seq_documentation_block:
ID   AAD13079 standard; DNA; 6703 BP.
XX
AC   AAD13079;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-2XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-2XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain  -
XX
PS   Example 4; Page 134-135; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-2XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domains-Mammalia caspase-3
CC   fusion protein.
XX
SQ   Sequence 6703 BP; 1653 A; 1677 C; 1674 G; 1699 T; 0 other;


alignment_scores:
            Quality:  275.00              Length:    275
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AAD13079   ..

 Align seg 1/1  to: AAD13079  from: 1  to: 6703

    3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1411 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAA 1460

   19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1461 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1510

   36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1511 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 1560

   53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1561 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 1610

   69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1611 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 1660

   86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1661 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1710

  103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1711 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 1760

  119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1761 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 1810

  136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1811 TGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 1860

  153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1861 GGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 1910

  169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1911 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 1960

  186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1961 AAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2010

  203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2011 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 2060

  219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2061 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 2110

  236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2111 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 2160
```

```
     253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2161 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 2210

     269 uThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||
    2211 CACAAAAGAACTCTATTTTTATCAC 2235

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13080

seq_documentation_block:
ID   AAD13080 standard; DNA; 6946 BP.
XX
AC   AAD13080;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-3XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-3XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain  -
XX
PS   Example 4; Page 136-137; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-3XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domains-Mammalia caspase-3
CC   fusion protein.
XX
SQ   Sequence 6946 BP; 1723 A; 1742 C; 1735 G; 1746 T; 0 other;

alignment_scores:
           Quality:  275.00            Length:    275
               Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AAD13080   ..

 Align seg 1/1  to: AAD13080  from: 1  to: 6946

       3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1654 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAA 1703

      19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1704 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1753

      36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1754 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 1803

      53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1804 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 1853

      69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1854 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 1903

      86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1904 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1953

     103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1954 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 2003

     119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2004 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 2053

     136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2054 TGAAAAAAATAACAAACTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 2103

     153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2104 GGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 2153

     169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2154 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 2203

     186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2204 AAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2253

     203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2254 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 2303

     219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2304 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 2353

     236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2354 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 2403

     253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2404 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 2453

     269 uThrLysGluLeuTyrPheTyrHis 277
```

```
          ||||||||||||||||||||||||
    2454 CACAAAAGAACTCTATTTTTATCAC 2478

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13081

seq_documentation_block:
ID   AAD13081 standard; DNA; 7189 BP.
XX
AC   AAD13081;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-4XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-4XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain  -
XX
PS   Example 4; Page 137-138; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-4XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domains-Mammalia caspase-3
CC   fusion protein.
XX
SQ   Sequence 7189 BP; 1793 A; 1807 C; 1796 G; 1793 T; 0 other;


alignment_scores:
            Quality:  275.00              Length:    275
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AAD13081   ..
```

```
Align seg 1/1  to: AAD13081  from: 1  to: 7189

       3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1897 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAA 1946

      19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1947 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1996

      36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1997 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 2046

      53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2047 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 2096

      69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2097 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 2146

      86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2147 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 2196

     103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2197 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 2246

     119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2247 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 2296

     136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2297 TGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 2346

     153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2347 GGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 2396

     169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2397 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 2446

     186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2447 AAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2496

     203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2497 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 2546

     219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2547 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 2596

     236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2597 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 2646

     253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2647 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 2696

     269 uThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||
    2697 CACAAAAGAACTCTATTTTTATCAC 2721

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81219

seq_documentation_block:
```

```
ID   AAX81219 standard; DNA; 835 BP.
XX
AC   AAX81219;
XX
DT   10-SEP-1999  (first entry)
XX
DE   Nucleotide sequence of caspase-3 (CPP32).
XX
KW   Rev-caspase; cysteine protease; zymogen; caspase; cancer; gene therapy;
KW   autoimmune disease; caspase-mediated apoptosis; neurodegenerative;
KW   tumour cell; myocardial infarction; human; ds.
XX
OS   Homo sapiens.
XX
PN   WO9935277-A2.
XX
PD   15-JUL-1999.
XX
PF   11-JAN-1999;   99WO-US00632.
XX
PR   09-JAN-1998;   98US-0070987.
XX
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES;
XX
DR   WPI; 1999-419353/35.
DR   P-PSDB; AAY21717.
XX
PT   New isolated nucleic acid molecule encoding a rev-caspase  - used
PT   for screening and identifying inhibitors or enhancers for treating
PT   cancer or autoimmune disease
XX
PS   Disclosure; Fig 13A-B; 74pp; English.
XX
CC   The invention relates to nucleic acid molecules encoding rev-caspases.
CC   Rev-caspases are cysteine proteases that specifically cleave proteins
CC   after Asp residues and is expressed as a zymogen, in which the small
CC   subunit is N-terminal to a large subunit. A gene delivery vehicle
CC   comprising a rev-caspase coding sequence is useful for the treatment of
CC   cancer, where the gene delivery vehicle is internalised by tumour cells.
CC   The gene delivery vehicle can also be used to treat autoimmune diseases.
CC   Cells transfected with a rev-caspase expressing vector can be used in
CC   identification of inhibitors or enhancers of caspase-mediated apoptosis.
CC   In vitro translated rev-caspase can be used to identify an inhibitor or
CC   enhancer of caspase processing activity. Caspase inhibitors are useful
CC   for treating neurodegenerative diseases as well as for inhibiting
CC   apoptosis in the heart following myocardial infarction. Sequences
CC   AAX81217 -AAX81226 represent human caspase genes encoding caspase 1-10
CC   gene products (AAY21715-Y21724).
XX
SQ   Sequence 835 BP; 274 A; 147 C; 178 G; 235 T; 1 other;

alignment_scores:
            Quality:  189.00              Length:    189
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AAX81219   ..

 Align seg 1/1  to: AAX81219  from: 1  to: 835

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 50

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 100

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 150

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 200

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 250

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 300

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 350

     117 lLeuLeuSerHisGlyGluGluGlyIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||
     351 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 400

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 450

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 500

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 550

     184 ysHisLysIleProVal 189
         ||||||||||||||||
     551 GTCATAAAATACCAGTG 567

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAS10503

seq_documentation_block:
ID   AAS10503 standard; DNA; 2635 BP.
XX
AC   AAS10503;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Human caspase 3 DNA sequence.
XX
KW   Human; caspase 3; apoptosis; hyperproliferative disorder; hepatitis;
KW   viral infection; haematopoietic disorder; autoimmune disorder;
KW   atherosclerosis; neurological disorder; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             225..1058
FT                   /*tag= a
FT                   /product= "Caspase 3"
XX
PN   WO200153310-A1.
XX
PD   26-JUL-2001.
XX
PF   11-JAN-2001; 2001WO-US00888.
XX
PR   18-JAN-2000; 2000US-0484617.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Zhang H,  Cowsert LM;
XX
```

```
DR   WPI; 2001-442252/47.
DR   P-PSDB; AAU05394.
XX
PT   New antisense compound to inhibit caspase 3 is useful for treating
PT   hepatitis and atherosclerosis -
XX
PS   Example 13; Page 92-94; 127pp; English.
XX
CC   The present sequence encoding for human caspase 3 is described
CC   in an invention relating to novel antisense oligonucleotides
CC   (AAS10517-AAS10676) and methods of using these compounds for the
CC   modulation of caspase 3 expression. The caspase 3 antisense
CC   oligonucleotides specifically hybridise with and inhibit the
CC   expression of caspase 3. Antisense compounds targeted to caspase 3
CC   are useful to inhibit caspase 3 expression in cells or tissues and
CC   to modulate apoptosis. The caspase 3 antisense oligonucleotides are
CC   useful for treating disorders associated with expression of caspase 3.
CC   Such disorders include hyperproliferative disorders (e.g. cancer),
CC   viral infections (e.g. hepatitis), haematopoietic disorders, autoimmune
CC   disorders, atherosclerosis and neurological disorders (e.g. Alzheimer's
CC   disease).
XX
SQ   Sequence 2635 BP; 815 A; 465 C; 569 G; 786 T; 0 other;


alignment_scores:
            Quality:  189.00              Length:    189
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AAS10503    ..

 Align seg 1/1  to: AAS10503  from: 1  to: 2635

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     225 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 274

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     275 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 324

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     325 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 424

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     425 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 474

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     475 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 524

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     525 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 574

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     575 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 624

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     625 TTGACCTGAAAAAAATAACAAACTTTTCAGAGGGGATCGTTGTAGAAGT 674

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||

     675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 724

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     725 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 774

     184 ysHisLysIleProVal 189
         ||||||||||||||||
     775 GTCATAAAATACCAGTG 791

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03910

seq_documentation_block:
ID   AAD03910 standard; cDNA; 2635 BP.
XX
AC   AAD03910;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Human caspase-3 cDNA.
XX
KW   Human; caspase-3; CPP32; yama; apopain;
KW   cysteine protease; apoptosis; caspase expression cassette; metastasis;
KW   tumour; cathepsin B; urokinase; proliferation; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             225..1058
FT                   /*tag= a
FT                   /product= "Human caspase-3"
XX
PN   WO200129232-A2.
XX
PD   26-APR-2001.
XX
PF   19-OCT-2000; 2000WO-US28941.
XX
PR   20-OCT-1999;   99US-0160559.
PR   14-AUG-2000; 2000US-0225564.
XX
PA   (SCIO-) SCIOS INC.
XX
PI   Cordell B,  Li Y;
XX
DR   WPI; 2001-290920/30.
DR   P-PSDB; AAE00600.
XX
PT   Novel fusion polypeptide comprising first and second caspase subunit
PT   separated by cleavage site not associated in nature with caspase
PT   subunit, useful for cloning gene encoding enzymes involved in
PT   proteolytic cleavage  -
XX
PS   Example 1; Fig 5; 116pp; English.
XX
CC   The present sequence is a cDNA encoding human Caspase-3 also known as
CC   CPP32, yama and apopain. Caspases are a family of cysteine proteases,
CC   that participate in the initiation and execution of apoptosis. Caspases
CC   exist as pro-enzymes, activated by cleavage into a large and small
CC   subunit, occurring after specific aspartic acid residues within the
CC   pro-enzyme sequence. The present invention relates to a method for
CC   functional cloning of genes encoding proteins or enzymes involved in
CC   proteolytic cleavage. The invention is based on the use of caspase
CC   expression cassettes comprising the coding sequence of a proteolytic
CC   cleavage site flanked by sequences encoding two caspase subunits. A
CC   fusion polypeptide comprising a first and a second caspase subunit,
CC   separated by a cleavage site not associated in nature, is useful for
CC   cloning gene encoding enzymes involved in proteolytic cleavage. An
CC   expression cassette containing fusion polypeptide is used to identify a
CC   mutant cell line deficient in an enzyme of interest and is also useful
CC   for diagnosis and suppression of proliferation or metastases of a tumour
CC   cell characterised by overexpression of a polypeptide (e.g. Cathepsin B
CC   or urokinase, selectively expressed in the tumour cells). DNA encoding
```

```
CC   fusion polypeptide is used in gene therapy.
XX
SQ   Sequence 2635 BP; 815 A; 465 C; 569 G; 786 T; 0 other;


alignment_scores:
            Quality:  189.00              Length:    189
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AAD03910    ..

 Align seg 1/1  to: AAD03910  from: 1  to: 2635

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     225 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 274

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     275 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 324

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     325 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 424

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     425 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 474

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     475 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 524

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     525 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 574

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     575 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 624

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     625 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 674

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 724

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     725 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 774

     184 ysHisLysIleProVal 189
         ||||||||||||||||
     775 GTCATAAAATACCAGTG 791

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13084

seq_documentation_block:
ID   AAD13084 standard; DNA; 7247 BP.
XX
AC   AAD13084;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-Ub-Met-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-Ub-Met-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain  -
XX
PS   Example 4; Page 140-141; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-Ub-Met-C3 construct comprising wildtype multimerised
CC   human ubiquitin destabilisation domain-Mammalia caspase-3 fusion protein.
XX
SQ   Sequence 7247 BP; 1794 A; 1796 C; 1826 G; 1831 T; 0 other;


alignment_scores:
            Quality:  180.00              Length:    180
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AAD13084    ..

 Align seg 1/1  to: AAD13084  from: 1  to: 7247

      10 SerLysSerIleLysAsnLeuGluProLysIleIleHisGlySerGluSe 26
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1171 TCAAAATCCATTAAAAATTGGAACCAAAGATCATACATGGAAGCGAATC 1220

      26 rMetAspSerGlyIleSerLeuAspAsnSerTyrLysMetAspTyrProG 43
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1221 AATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTG 1270

      43 luMetGlyLeuCysIleIleIleAsnAsnLysAsnPheHisLysSerThr 59
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1271 AGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACT 1320
```

```
     60 GlyMetThrSerArgSerGlyThrAspValAspAlaAlaAsnLeuArgGl 76
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1321 GGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGA 1370

     76 uThrPheArgAsnLeuLysTyrGluValArgAsnLysAsnAspLeuThrA 93
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1371 AACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACAC 1420

     93 rgGluGluIleValGluLeuMetArgAspValSerLysGluAspHisSer 109
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1421 GTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGC 1470

    110 LysArgSerSerPheValCysValLeuLeuSerHisGlyGluGluGlyIl 126
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1471 AAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAAT 1520

    126 eIlePheGlyThrAsnGlyProValAspLeuLysLysIleThrAsnPheP 143
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1521 AATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTT 1570

    143 heArgGlyAspArgCysArgSerLeuThrGlyLysProLysLeuPheIle 159
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1571 TCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATT 1620

    160 IleGlnAlaCysArgGlyThrGluLeuAspCysGlyIleGluThrAspSe 176
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1621 ATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAG 1670

    176 rGlyValAspAspAspMetAlaCysHisLysIleProVal 189
        |||||||||||||||||||||||||||||||||||||
   1671 TGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTG 1710
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV32608

seq_documentation_block:

```
ID   AAV32608 standard; cDNA; 1001 BP.
XX
AC   AAV32608;
XX
DT   26-OCT-1998  (first entry)
XX
DE   Mutant human apopain (caspase-3) pro-enzyme cDNA.
XX
KW   Mutant human apopain pro-enzyme; caspase-3; protease; SPA;
KW   scintillation proximity binding assay; diabetes; cancer;
KW   4-phosphono(difluoromethyl) phenylalanine group; osteoporosis;
KW   tyrosine phosphatase; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             64..897
FT                   /*tag= a
FT                   /product= "Mutant apopain C163S pro-enzyme"
FT   misc_feature    148..588
FT                   /*tag= b
FT                   /note= "Encodes the P17 subunit"
FT   mutation        550..552
FT                   /*tag= c
FT                   /note= "Changed from TGC in wild-type to TCC in
FT                    mutant"
FT   misc_feature    589..894
FT                   /*tag= d
FT                   /note= "Encodes the P12 subunit"
XX
PN   WO9820024-A1.
XX
PD   14-MAY-1998.
XX
PF   03-NOV-1997;   97WO-CA00824.
XX
PR   04-NOV-1996;   96US-0030411.
XX
PA   (MERI ) MERCK FROSST CANADA INC.
XX
PI   Desmarais S,  Friesen R,  Zamboni R;
XX
DR   WPI; 1998-348101/30.
DR   P-PSDB; AAW48937.
XX
PT   Peptide(s) useful in binding assays for tyrosine phosphatases or
PT   cysteine proteases - contain two or more 4-phosphono(difluoromethyl)
PT   phenylalanine groups to improve binding affinity
XX
PS   Disclosure; Pages 38-39; 59pp; English.
XX
CC   The present sequence represents a mutated human apopain (caspase-3)
CC   pro-enzyme cDNA.  The cDNA encodes a mutant human apopain C163S
CC   pro-enzyme.  The invention provides a method for use in a
CC   scintillation proximity binding assay (SPA) for proteases and
CC   phosphatases.  The method involves using mutated proteases and
CC   phosphatases whereby the catalytic cysteine residue of the enzymes
CC   are replaced with a serine or alanine residue to correct the problem of
CC   interference in SPA from extraneous oxidising and alkylating agents.
CC   The mutation affects the catalytic properties of the enzyme but does
CC   not affect their binding properties.  The invention claims for new
CC   ligands for use in SPA which have increased binding affinity for a
CC   tyrosine phosphatase or cysteine protease.  The ligands contain at least
CC   two 4-phosphono(difluoromethyl) phenylalanine groups which increase
CC   binding affinity of the ligand to its respective enzyme.  The assay can
CC   be used to determine the ability of new ligands and compound mixtures to
CC   competitively bind with an enzyme.  The method is claimed to allow
CC   a better usage of SPA in the discovery of compounds for the treatment
CC   and study of diseases, e.g. diabetes, cancer and osteoporosis.
XX
SQ   Sequence 1001 BP; 310 A; 172 C; 226 G; 293 T; 0 other;
```

alignment_scores:

```
           Quality:  176.00              Length:    277
             Ratio:  0.638                 Gaps:      0
 Percent Similarity:  99.639   Percent Identity: 99.639
```

alignment_block:

US-09-613-508-4 x AAV32608 ..

Align seg 1/1 to: AAV32608 from: 1 to: 1001

```
      1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

     17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

     34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

     51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413
```

```
     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         |||||||||||||||||||||||||||||||||   ||||||||||||
     514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTCCCGTGGTACAGA 563

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         |||||||||||||||||||||||||||||||
     864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV32615

seq_documentation_block:
ID   AAV32615 standard; cDNA; 1001 BP.
XX
AC   AAV32615;
XX
DT   26-OCT-1998  (first entry)
XX
DE   Mutant human apopain (caspase-3) pro-enzyme cDNA.
XX
KW   Mutant human apopain pro-enzyme; caspase-3; protease; SPA;
KW   scintillation proximity binding assay; diabetes; cancer;
KW   4-phosphono(difluoromethyl) phenylalanine group; osteoporosis;
KW   tyrosine phosphatase; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             64..897
FT                   /*tag=  a
FT                   /product= "Mutant apopain C163S pro-enzyme"
FT   misc_feature    148..588
FT                   /*tag=  b
FT                   /note= "Encodes the P17 subunit"
FT   mutation        550..552
FT                   /*tag=  c
FT                   /note= "Changed from TGC in wild-type to TCC in
FT                        mutant"
FT   misc_feature    589..894
FT                   /*tag=  d
FT                   /note= "Encodes the P12 subunit"
XX
PN   WO9820156-A1.
XX
PD   14-MAY-1998.
XX
PF   03-NOV-1997;   97WO-CA00825.
XX
PR   04-NOV-1996;   96US-0030408.
XX
PA   (MERI ) MERCK FROSST CANADA INC.
XX
PI   Desmarais S,  Ford-Hutchinson A,  Friesen R,  Gresser M;
PI   Kennedy B,  Nicholson D,  Ramachandran C,  Skorey K;
XX
DR   WPI; 1998-348103/30.
XX
PT   Enzyme binding assay for detection of useful compounds - uses mutant
PT   form of wild-type enzyme, in which serine replaces cysteine at
PT   active site, to reduce interference from oxidlsing/alkylating agents
XX
PS   Disclosure; Pages 38-39; 63pp; English.
XX
CC   The present sequence represents a mutated human apopain (caspase-3)
CC   pro-enzyme cDNA.  The cDNA encodes a mutant human apopain C163S
CC   pro-enzyme.  The invention claims for a method for use in a
CC   scintillation proximity binding assay (SPA) for proteases and
CC   phosphatases.  The method involves using mutated proteases and
CC   phosphatases whereby the catalytic cysteine residue of the enzymes
CC   are replaced with a serine or alanine residue to correct the problem of
CC   interference in SPA from extraneous oxidising and alkylating agents.
CC   The mutation affects the catalytic properties of the enzyme but does
CC   not affect their binding properties.  The invention provides new
CC   ligands for use in SPA which have increased binding affinity for a
CC   tyrosine phosphatase or cysteine protease.  The ligands contain at least
CC   two 4-phosphono(difluoromethyl) phenylalanine groups which increase
CC   binding affinity of the ligand to its respective enzyme.  The assay can
CC   be used to determine the ability of new ligands and compound mixtures to
CC   competitively bind with an enzyme.  The method is claimed to allow
CC   a better usage of SPA in the discovery of compounds for the treatment
CC   and study of diseases, e.g. diabetes, cancer and osteoporosis.
XX
SQ   Sequence 1001 BP; 310 A; 172 C; 226 G; 293 T; 0 other;


alignment_scores:
           Quality:  176.00              Length:    277
             Ratio:  0.638                 Gaps:      0
 Percent Similarity:  99.639   Percent Identity: 99.639

alignment_block:
 US-09-613-508-4 x AAV32615    ..

 Align seg 1/1  to: AAV32615  from: 1  to: 1001

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

 84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
    ||||||||||||||||||||||||||||||||||||||||||||||||||
314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
    ||||||||||||||||||||||||||||||||||||||||||||||||||
364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
    ||||||||||||||||||||||||||||||||||||||||||||||||||
414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
    |||||||||||||||||||||||||||||||||  ||||||||||||
514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTCCCGTGGTACAGA 563

167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
    ||||||||||||||||||||||||||||||||||||||||||||||||||
564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
    ||||||||||||||||||||||||||||||||||||||||||||||||||
664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
    ||||||||||||||||||||||||||||||||||||||||||||||||||
714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
    ||||||||||||||||||||||||||||||||||||||||||||||||||
764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
    ||||||||||||||||||||||||||||||||||||||||||||||||||
814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
    ||||||||||||||||||||||||||||||
864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894
```

```
OM of: US-09-613-508-4 to: GenEmbl:*   out_format : pfs

Date: Aug 15, 2002 10:16 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO_spool/US09613508/runat_14082002_112714_7772/app_query.fasta_1.701
 -DB=GenEmbl -QFMT=fastap -SUFFIX=olig.rge -GAPOP=4.500
 -GAPEXT=0.050 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000 -YGAPEXT=60.000
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=oligo
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality
 -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext
 -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US09613508_@CGN1_1_4651 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-613-508-4
Query length: 277
Database: GenEmbl:*
Database sequences: 1797656
Database length: 1873333701
Search time (sec): 2733.790000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000


score_list:
Sequence              Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_pat:AX207737         +  277.00 5533.74  7.2e-300  834  ! AX207737 Sequence 35 from Paten
gb_pr:HSU26943          +  277.00 5533.74  7.2e-300  834  ! U26943 Human cysteine protease
gb_pat:AR109589         +  277.00 5532.44  8.5e-300 1001  ! AR109589 Sequence 5 from paten
gb_pr:HSU13738          +  277.00 5532.44  8.5e-300 1001  ! U13738 Human cysteine protease
gb_pat:AR103059         +  277.00 5531.39  9.7e-300 1159  ! AR103059 Sequence 3 from paten
gb_pr:BC016926          +  277.00 5525.46  2.1e-299 2652  ! BC016926 Homo sapiens, clone M
gb_pat:AX207740         +  275.00 5478.97  8.1e-297 6436  ! AX207740 Sequence 38 from Pate
gb_pat:AX207741         +  275.00 5478.68  8.4e-297 6703  ! AX207741 Sequence 39 from Pate
gb_pat:AX207742         +  275.00 5478.43  8.7e-297 6946  ! AX207742 Sequence 40 from Pate
gb_pat:AX207743         +  275.00 5478.18  9.0e-297 7189  ! AX207743 Sequence 41 from Pate
gb_pat:AR172878         +  189.00 3758.98  5.2e-201 2635  ! AR172878 Sequence 3 from paten
gb_pat:AX134409         +  189.00 3758.98  5.2e-201 2635  ! AX134409 Sequence 5 from Paten
gb_pat:AX331135         +  189.00 3758.98  5.2e-201 2635  ! AX331135 Sequence 1644 from Pa
gb_pr:HSU13737          +  189.00 3758.98  5.2e-201 2635  ! U13737 Human cysteine protease
gb_pat:AX207746         +  180.00 3571.07  1.5e-190 7248  ! AX207746 Sequence 44 from Pate
gb_pat:AR109591         +  176.00 3504.93  7.3e-187 1001  ! AR109591 Sequence 9 from paten
gb_sy:AF052645          +  175.00 3485.84  8.4e-186  873  ! AF052645 Synthetic construct re
gb_pat:AX260262         +  167.00 3320.59  1.4e-176 1673  ! AX260262 Sequence 27 from Pate
gb_pat:AX260259         +  167.00 3317.53  2.0e-176 2566  ! AX260259 Sequence 24 from Pate
gb_pr:HSA413269         +  153.00 3044.53  3.2e-161  834  ! AJ413269 Homo sapiens mRNA for
gb_sy:AF052647          +  147.00 2923.88  1.7e-154  858  ! AF052647 Synthetic construct re
gb_pat:AX134427         +  146.00 2904.44  2.0e-153  786  ! AX134427 Sequence 23 from Paten
gb_pat:AX235203         +  146.00 2893.52  8.3e-153 3614  ! AX235203 Sequence 1 from Paten
gb_htg:AC084049         -   76.00 1460.77  5.3e-73  169950  ! AC084049 Homo sapiens chromo
gb_htg:AC079257         -   76.00 1460.43  5.5e-73  178290  ! AC079257 Homo sapiens chromo
gb_htg:AC060790         +   76.00 1459.35  6.4e-73  207425  ! AC060790 Homo sapiens chromo
gb_om:AB029345          +   36.00  695.84  2.1e-30  834  ! AB029345 Sus scrofa mRNA for ca
gb_ro:D86352            +   35.00  675.43  2.9e-29  874  ! D86352 Mouse mRNA for CPP32, co
gb_pat:AR172885         +   35.00  672.61  4.2e-29 1297  ! AR172885 Sequence 10 from pate
gb_ro:MMCASP3           +   35.00  672.61  4.2e-29 1297  ! Y13086 M.musculus mRNA for cas
gb_ro:MMU19522          +   35.00  671.92  4.6e-29 1429  ! U19522 Mus musculus cysteine p
gb_ro:MMU49929          +   35.00  671.73  4.7e-29 1466  ! U49929 Mus musculus ICE-like c
gb_ro:MMU63720          +   35.00  671.07  5.1e-29 1609  ! U63720 Mus musculus CPP32 apop
gb_om:AF068837          +   30.00  577.02  8.9e-24  665  ! AF068837 Ovis aries caspase-3 m
gb_ro:RNU58656          +   30.00  574.98  1.2e-23  884  ! U58656 Rattus norvegicus cystei
gb_ro:RNU84410          +   30.00  573.54  1.4e-23 1081  ! U84410 Rattus norvegicus inter
gb_pat:AR059697         +   30.00  572.93  1.5e-23 1177  ! AR059697 Sequence 1 from paten
gb_pat:AR059698         +   30.00  572.93  1.5e-23 1177  ! AR059698 Sequence 3 from paten
gb_ro:RNU49930          +   30.00  572.51  1.6e-23 1249  ! U49930 Rattus norvegicus I
gb_ro:RNU34685          +   26.00  497.01  2.5e-19  639  ! U34685 Rattus norvegicus cy
gb_ro:CGU27463          +   25.00  467.48  1.1e-17 2395  ! U27463 Cricetinae gen. sp.
gb_htg:AC096150         +   25.00  440.45  3.6e-16 104747  ! AC096150 Rattus norvegic
gb_ov:AF083029          +   23.00  430.85  1.2e-15 1464  ! AF083029 Gallus gallus cas
gb_ro:MMLICE03          +   23.00  420.08  4.9e-15 6598  ! U54803 Mus musculus cystei
gb_pr:AL592104          -   23.00  398.67  7.6e-14 131489  ! AL592104 Human DNA seque

seq_name: gb_pat:AX207737

seq_documentation_block:
LOCUS       AX207737                834 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 35 from Patent WO0157242.
ACCESSION   AX207737
VERSION     AX207737.1  GI:15422421
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 834)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 35 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .834
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      274 a    147 c    179 g    234 t
ORIGIN


alignment_scores:
           Quality:  277.00              Length:    277
             Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AX207737   ..

 Align seg 1/1  to: AX207737  from: 1  to: 834

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 50

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 100

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 150

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 200

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 250

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 300

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 350

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     351 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 400

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 450

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 500

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 550

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     551 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 600

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     601 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 650

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     651 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 700

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     701 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 750

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     751 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 800

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||
     801 CATGCTCACAAAAGAACTCTATTTTTATCAC 831

seq_name: gb_pr:HSU26943

seq_documentation_block:
LOCUS       HSU26943                834 bp    mRNA    linear   PRI 29-AUG-1995
DEFINITION  Human cysteine protease Yama mRNA, complete cds.
ACCESSION   U26943
VERSION     U26943.1  GI:857568
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 834)
  AUTHORS   Tewari,M., Quan,L.T., O'Rourke,K., Desnoyers,S., Zeng,Z.,
            Beidler,D.R., Poirier,G.G., Salvesen,G.S. and Dixit,V.M.
  TITLE     Yama/CPP32 beta, a mammalian homolog of CED-3, is a
            CrmA-inhibitable protease that cleaves the death substrate
            poly(ADP-ribose) polymerase
  JOURNAL   Cell 81 (5), 801-809 (1995)
  MEDLINE   95292347
REFERENCE   2  (bases 1 to 834)
  AUTHORS   Tewari,M., O'Rourke,K., Lin,Z. and Dixit,V.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-1995) Muneesh Tewari, Pathology, University of
            Michigan, 1301 Catherine, Ann Arbor, MI 48109, USA
FEATURES             Location/Qualifiers
     source          1. .834
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="umbilical vein endothelial cells treated with
                     tumor necrosis factor and cycloheximide"
     CDS             1. .834
                     /note="cysteine protease CPP32beta; CED-3/ICE protein
                     family"
                     /codon_start=1
                     /product="Yama protein"
                     /protein_id="AAA74929.1"
                     /db_xref="GI:857569"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVEADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
BASE COUNT      274 a    147 c    179 g    234 t
ORIGIN


alignment_scores:
           Quality:  277.00            Length:    277
             Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x HSU26943   ..

 Align seg 1/1  to: HSU26943  from: 1  to: 834

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 50

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 100

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 150

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 200

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 250

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 300

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 350

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 400

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 450

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 500

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 550

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     551 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 600

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     601 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 650

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     651 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 700

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     701 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 750

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     751 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 800

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
.    801 CATGCTCACAAAAGAACTCTATTTTTATCAC 831

seq_name: gb_pat:AR109589

seq_documentation_block:
LOCUS       AR109589              1001 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 5 from patent US 6114132.
ACCESSION   AR109589
VERSION     AR109589.1  GI:12825865
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1001)
  AUTHORS   Desmarais,S., Gresser,M., Friesen,R., Kennedy,B., Nicholson,D.,
            Skorey,K., Ramachandran,C. and Ford-Hutchinson,A.
  TITLE     Prosphatase binding assay
  JOURNAL   Patent: US 6114132-A 5 05-SEP-2000;
FEATURES             Location/Qualifiers
     source          1. .1001
                     /organism="unknown"
BASE COUNT      310 a    171 c    227 g    293 t
ORIGIN


alignment_scores:
            Quality:  277.00             Length:    277
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AR109589   ..

 Align seg 1/1  to: AR109589  from: 1  to: 1001

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 563

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: gb_pr:HSU13738

seq_documentation_block:
LOCUS       HSU13738              1001 bp    mRNA    linear   PRI 14-APR-1995
DEFINITION  Human cysteine protease CPP32 isoform beta mRNA, complete cds.
ACCESSION   U13738
VERSION     U13738.1  GI:561667
KEYWORDS    cysteine protease; interleukin 1-beta converting enzyme; apoptotic
            protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1001)
  AUTHORS   Fernandes-Alnemri,T., Litwack,G. and Alnemri,E.S.
  TITLE     CPP32, a novel human apoptotic protein with homology to
            Caenorhabditis elegans cell death protein Ced-3 and mammalian
            interleukin-1 beta-converting enzyme
  JOURNAL   J. Biol. Chem. 269 (49), 30761-30764 (1994)
  MEDLINE   95074098
REFERENCE   2  (bases 1 to 1001)
  AUTHORS   Fernandes-Alnemri,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-AUG-1994) Teresa Fernandes-Alnemri, Pharmacology,
            Thomas Jefferson University, Jefferson Cancer Institute, 233, S.
            Tenth Street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
```

```
     source          1. .1001
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="J6.1 Pst I fragment"
                     /cell_line="Jurkat"
                     /cell_type="T-lymphocyte"
                     /clone_lib="Human Jurkat T-lymphocyte lambda Uni-ZAP XR
                     cDNA library"
     5'UTR           1. .63
                     /note="An alternatively spliced isoform which lacks most
                     of the 5' nontranslated sequence"
     CDS             64. .897
                     /note="new member of the interleukin 1-beta converting
                     enzyme gene family of cysteine proteases"
                     /codon_start=1
                     /product="cysteine protease CPP32 isoform beta"
                     /protein_id="AAB60355.1"
                     /db_xref="GI:561668"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVEADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
     misc_difference 234
                     /note="differs from isoform alpha by one bp substitution
                     at position 234"
                     /replace="a"
     misc_difference 633
                     /note="differs from isoform alpha by one bp substitution
                     at position 633"
                     /replace="t"
BASE COUNT      310 a    171 c    227 g    293 t
ORIGIN


alignment_scores:
            Quality:  277.00             Length:    277
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x HSU13738   ..

 Align seg 1/1  to: HSU13738  from: 1  to: 1001

      1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

     17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

     34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

     51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 563

    167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

    184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

    201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

    217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

    234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

    251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

    267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
        ||||||||||||||||||||||||||||||
    864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: gb_pat:AR103059

seq_documentation_block:
LOCUS       AR103059               1159 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 3 from patent US 6087150.
ACCESSION   AR103059
VERSION     AR103059.1  GI:12814647
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1159)
  AUTHORS   He,W.W., Rosen,C.A., Hudson,P.L. and Hastings,G.A.
  TITLE     Interleukin-1 .beta. converting enzyme like apoptosis protease-3
            and 4
  JOURNAL   Patent: US 6087150-A 3 11-JUL-2000;
FEATURES             Location/Qualifiers
     source          1. .1159
                     /organism="unknown"
BASE COUNT      329 a    220 c    276 g    334 t
ORIGIN


alignment_scores:
            Quality:  277.00             Length:    277
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AR103059   ..

 Align seg 1/1  to: AR103059  from: 1  to: 1159
```

```
     1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   223 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 272

    17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   273 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 322

    34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   323 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 372

    51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   373 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 422

    67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   423 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 472

    84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   473 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 522

   101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   523 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 572

   117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   573 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 622

   134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   623 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 672

   151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   673 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 722

   167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   723 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 772

   184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   773 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 822

   201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   823 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 872

   217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   873 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 922

   234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   923 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 972

   251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   973 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1022

   267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
       ||||||||||||||||||||||||||||||
  1023 CATGCTCACAAAAGAACTCTATTTTTATCAC 1053

seq_name: gb_pr:BC016926

seq_documentation_block:
LOCUS       BC016926               2652 bp    mRNA    linear   PRI 09-NOV-2001
DEFINITION  Homo sapiens, clone MGC:21241 IMAGE:4419175, mRNA, complete cds.
ACCESSION   BC016926
VERSION     BC016926.1  GI:16877343
KEYWORDS    MGC.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2652)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 27 Row: j Column: 23
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 14790118.
FEATURES             Location/Qualifiers
     source          1. .2652
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="MGC:21241 IMAGE:4419175"
                     /tissue_type="Lymph, lymphoma"
                     /clone_lib="NIH_MGC_85"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             228. .1061
                     /codon_start=1
                     /product="Unknown (protein for MGC:21241)"
                     /protein_id="AAH16926.1"
                     /db_xref="GI:16877344"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVEADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
BASE COUNT      824 a    466 c    576 g    786 t
ORIGIN


alignment_scores:
            Quality:  277.00              Length:    277
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x BC016926   ..

 Align seg 1/1  to: BC016926  from: 1  to: 2652

      1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    228 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 277

     17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
    278 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 327

     34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    328 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 377

     51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    378 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 427

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    428 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 477

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    478 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 527

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    528 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 577

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    578 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 627

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    628 TTGACCTGAAAAAAATAACAAACTTTTCAGAGGGGATCGTTGTAGAAGT 677

    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    678 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 727

    167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    728 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 777

    184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    778 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 827

    201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    828 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 877

    217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    878 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 927

    234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    928 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 977

    251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    978 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1027

    267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
        ||||||||||||||||||||||||||||||
   1028 CATGCTCACAAAAGAACTCTATTTTTATCAC 1058

seq_name: gb_pat:AX207740

seq_documentation_block:
LOCUS       AX207740               6436 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 38 from Patent WO0157242.
ACCESSION   AX207740
VERSION     AX207740.1  GI:15422424
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.

REFERENCE   1  (bases 1 to 6436)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 38 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6436
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-1XUb-C3 construct"
BASE COUNT     1577 a   1606 c   1607 g   1646 t
ORIGIN


alignment_scores:
            Quality:  275.00              Length:    275
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AX207740   ..

 Align seg 1/1  to: AX207740  from: 1  to: 6436

      3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1144 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAA 1193

     19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1194 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1243

     36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1244 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 1293

     53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1294 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 1343

     69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1344 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 1393

     86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1394 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1443

    103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1444 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 1493

    119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1494 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 1543

    136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1544 TGAAAAAAATAACAAACTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 1593

    153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1594 GGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 1643

    169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1644 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 1693

    186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1694 AAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 1743

    203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    1744 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 1793

     219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    1794 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 1843

     236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    1844 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 1893

     253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    1894 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 1943

     269 uThrLysGluLeuTyrPheTyrHis 277
          ||||||||||||||||||||||||
    1944 CACAAAAGAACTCTATTTTTATCAC 1968

seq_name: gb_pat:AX207741

seq_documentation_block:
LOCUS       AX207741               6703 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 39 from Patent WO0157242.
ACCESSION   AX207741
VERSION     AX207741.1  GI:15422425
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 6703)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 39 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6703
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-2XUb-C3 construct"
BASE COUNT     1653 a   1677 c   1674 g   1699 t
ORIGIN


alignment_scores:
            Quality:  275.00            Length:    275
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AX207741   ..

 Align seg 1/1  to: AX207741  from: 1  to: 6703

       3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1411 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAA 1460

      19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1461 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1510

      36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1511 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 1560

      53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1561 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 1610

      69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1611 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 1660

      86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1661 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1710

     103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1711 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 1760

     119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1761 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 1810

     136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1811 TGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 1860

     153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1861 GGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 1910

     169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1911 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 1960

     186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1961 AAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2010

     203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2011 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 2060

     219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2061 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 2110

     236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2111 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 2160

     253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2161 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 2210

     269 uThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||
    2211 CACAAAAGAACTCTATTTTTATCAC 2235

seq_name: gb_pat:AX207742

seq_documentation_block:
LOCUS       AX207742               6946 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 40 from Patent WO0157242.
ACCESSION   AX207742
VERSION     AX207742.1  GI:15422426
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 6946)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 40 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6946
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-3XUb-C3 construct"
BASE COUNT     1723 a   1742 c   1735 g   1746 t
ORIGIN
```

```
alignment_scores:
           Quality:  275.00              Length:    275
             Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AX207742   ..

 Align seg 1/1  to: AX207742  from: 1  to: 6946

       3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1654 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAA 1703

      19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1704 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1753

      36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1754 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 1803

      53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1804 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 1853

      69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1854 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 1903

      86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1904 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1953

     103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1954 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 2003

     119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2004 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 2053

     136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2054 TGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 2103

     153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2104 GGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 2153

     169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2154 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 2203

     186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2204 AAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2253

     203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2254 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 2303

     219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2304 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 2353

     236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2354 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 2403

     253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2404 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 2453

     269 uThrLysGluLeuTyrPheTyrHis 277
         |||||||||||||||||||||||||
    2454 CACAAAAGAACTCTATTTTTATCAC 2478

seq_name: gb_pat:AX207743

seq_documentation_block:
LOCUS       AX207743               7189 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 41 from Patent WO0157242.
ACCESSION   AX207743
VERSION     AX207743.1  GI:15422427
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 7189)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 41 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7189
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-4XUb-C3 construct"
BASE COUNT     1793 a   1807 c   1796 g   1793 t
ORIGIN


alignment_scores:
           Quality:  275.00              Length:    275
             Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AX207743   ..

 Align seg 1/1  to: AX207743  from: 1  to: 7189

       3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1897 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAA 1946

      19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1947 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1996

      36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1997 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 2046

      53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2047 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 2096

      69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2097 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 2146

      86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2147 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 2196

     103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2197 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 2246

     119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2247 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 2296
```

```
     136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2297 TGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 2346

     153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2347 GGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 2396

     169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2397 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 2446

     186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2447 AAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2496

     203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2497 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 2546

     219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2547 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 2596

     236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2597 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 2646

     253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2647 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 2696

     269 uThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||
    2697 CACAAAAGAACTCTATTTTTATCAC 2721
```

seq_name: gb_pat:AR172878

seq_documentation_block:
LOCUS AR172878 2635 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 3 from patent US 6303374.
ACCESSION AR172878
VERSION AR172878.1 GI:17912369
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 2635)
AUTHORS Zhang,H. and Cowsert,L.M.
TITLE Antisense modulation of caspase 3 expression
JOURNAL Patent: US 6303374-A 3 16-OCT-2001;
FEATURES Location/Qualifiers
source 1. .2635
/organism="unknown"
BASE COUNT 815 a 465 c 569 g 786 t
ORIGIN

alignment_scores:
Quality: 189.00 Length: 189
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-613-508-4 x AR172878 ..

Align seg 1/1 to: AR172878 from: 1 to: 2635

```
       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     225 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 274

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     275 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 324

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     325 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 424

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     425 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 474

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     475 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 524

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     525 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 574

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     575 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 624

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     625 TTGACCTGAAAAAAATAACAAACTTTTCAGAGGGGATCGTTGTAGAAGT 674

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 724

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     725 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 774

     184 ysHisLysIleProVal 189
         |||||||||||||||||
     775 GTCATAAAATACCAGTG 791
```

seq_name: gb_pat:AX134409

seq_documentation_block:
LOCUS AX134409 2635 bp DNA linear PAT 29-MAY-2001
DEFINITION Sequence 5 from Patent WO0129232.
ACCESSION AX134409
VERSION AX134409.1 GI:14271055
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2635)
AUTHORS Cordell,B. and Li,Y.
TITLE Functional cloning of genes encoding proteins/enzymes involved in proteolytic cleavage
JOURNAL Patent: WO 0129232-A 5 26-APR-2001;
Scios Inc. (US)
FEATURES Location/Qualifiers
source 1. .2635
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS 225. .1058
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC39521.1"
/db_xref="GI:14271056"
/translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII

```
                    QACRGTELDCGIETDSGVDDDMACHKIPVDADFLYAYSTAPGYYSWRNSKDGSWFIQS
                    LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                    H"
BASE COUNT      815 a    465 c    569 g    786 t
ORIGIN


alignment_scores:
            Quality:  189.00              Length:    189
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AX134409   ..

 Align seg 1/1  to: AX134409  from: 1  to: 2635

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     225 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 274

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     275 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 324

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     325 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 424

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     425 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 474

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     475 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 524

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     525 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 574

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     575 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 624

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     625 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 674

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 724

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     725 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 774

     184 ysHisLysIleProVal 189
         ||||||||||||||||
     775 GTCATAAAATACCAGTG 791

seq_name: gb_pat:AX331135

seq_documentation_block:
LOCUS       AX331135               2635 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 1644 from Patent WO0194629.
ACCESSION   AX331135
VERSION     AX331135.1  GI:18121769
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 1644 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .2635
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      815 a    465 c    569 g    786 t
ORIGIN


alignment_scores:
            Quality:  189.00              Length:    189
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AX331135   ..

 Align seg 1/1  to: AX331135  from: 1  to: 2635

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     225 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 274

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     275 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 324

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     325 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 424

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     425 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 474

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     475 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 524

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     525 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 574

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     575 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 624

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     625 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 674

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 724

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     725 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 774
```

```
     184 ysHisLysIleProVal 189
         ||||||||||||||||
     775 GTCATAAAATACCAGTG 791

seq_name: gb_pr:HSU13737

seq_documentation_block:
LOCUS       HSU13737                2635 bp    mRNA    linear   PRI 14-APR-1995
DEFINITION  Human cysteine protease CPP32 isoform alpha mRNA, complete cds.
ACCESSION   U13737
VERSION     U13737.1  GI:561665
KEYWORDS    cysteine protease; interleukin 1-beta converting enzyme; apoptotic
            protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2635)
  AUTHORS   Fernandes-Alnemri,T., Litwack,G. and Alnemri,E.S.
  TITLE     CPP32, a novel human apoptotic protein with homology to
            Caenorhabditis elegans cell death protein Ced-3 and mammalian
            interleukin-1 beta-converting enzyme
  JOURNAL   J. Biol. Chem. 269 (49), 30761-30764 (1994)
  MEDLINE   95074098
REFERENCE   2  (bases 1 to 2635)
  AUTHORS   Fernandes-Alnemri,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-AUG-1994) Teresa Fernandes-Alnemri, Pharmacology,
            Thomas Jefferson University, Jefferson Cancer Institute, 233, S.
            Tenth Street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .2635
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="J1.1 EcoRI/XhoI fragment"
                     /cell_line="Jurkat"
                     /cell_type="T-lymphocyte"
                     /clone_lib="Human Jurkat T-lymphocyte lambda Uni-ZAP XR
                     cDNA library"
     CDS             225. .1058
                     /note="new member of the interleukin 1-beta converting
                     enzyme gene family of cysteine proteases"
                     /codon_start=1
                     /product="cysteine protease CPP32 isoform alpha"
                     /protein_id="AAA65015.1"
                     /db_xref="GI:561666"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVDADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
     polyA_site      2635
                     /note="19 A residues"
BASE COUNT      815 a    465 c    569 g    786 t
ORIGIN


alignment_scores:
            Quality:  189.00            Length:    189
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x HSU13737   ..

 Align seg 1/1  to: HSU13737  from: 1  to: 2635

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     225 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 274

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     275 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 324

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     325 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 424

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     425 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 474

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     475 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 524

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     525 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 574

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     575 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 624

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     625 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 674

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 724

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     725 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 774

     184 ysHisLysIleProVal 189
         ||||||||||||||||
     775 GTCATAAAATACCAGTG 791

seq_name: gb_pat:AX207746

seq_documentation_block:
LOCUS       AX207746                7248 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 44 from Patent WO0157242.
ACCESSION   AX207746
VERSION     AX207746.1  GI:15422430
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 7248)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 44 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7248
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-Ub-Met-C3 construct"
BASE COUNT     1794 a   1796 c   1826 g   1832 t
ORIGIN


alignment_scores:
            Quality:  180.00            Length:    180
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000
```

```
alignment_block:
 US-09-613-508-4 x AX207746   ..

 Align seg 1/1  to: AX207746  from: 1  to: 7248

      10 SerLysSerIleLysAsnLeuGluProLysIleIleHisGlySerGluSe 26
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1171 TCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATC 1220

      26 rMetAspSerGlyIleSerLeuAspAsnSerTyrLysMetAspTyrProG 43
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1221 AATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTG 1270

      43 luMetGlyLeuCysIleIleIleAsnAsnLysAsnPheHisLysSerThr 59
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1271 AGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACT 1320

      60 GlyMetThrSerArgSerGlyThrAspValAspAlaAlaAsnLeuArgGl 76
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1321 GGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGA 1370

      76 uThrPheArgAsnLeuLysTyrGluValArgAsnLysAsnAspLeuThrA 93
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1371 AACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACAC 1420

      93 rgGluGluIleValGluLeuMetArgAspValSerLysGluAspHisSer 109
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1421 GTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGC 1470

     110 LysArgSerSerPheValCysValLeuLeuSerHisGlyGluGluGlyIl 126
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1471 AAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAAT 1520

     126 eIlePheGlyThrAsnGlyProValAspLeuLysLysIleThrAsnPheP 143
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1521 AATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTT 1570

     143 heArgGlyAspArgCysArgSerLeuThrGlyLysProLysLeuPheIle 159
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1571 TCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATT 1620

     160 IleGlnAlaCysArgGlyThrGluLeuAspCysGlyIleGluThrAspSe 176
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1621 ATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAG 1670

     176 rGlyValAspAspAspMetAlaCysHisLysIleProVal 189
         |||||||||||||||||||||||||||||||||||||
    1671 TGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTG 1710
```

```
OM of: US-09-613-508-4 to: EST:*   out_format : pfs

Date: Aug 15, 2002  9:29 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO_spool/US09613508/runat_14082002_112713_7761/app_query.fasta_1.701
 -DB=EST -QFMT=fastap -SUFFIX=olig.rst -GAPOP=4.500 -GAPEXT=0.050
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=60.000 -YGAPEXT=60.000 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15
 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500 -MINLEN=0
 -MAXLEN=2000000000 -USER=US09613508_@CGN1_1_3002 -NCPU=6
 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-09-613-508-4
Query length: 277
Database: EST:*
Database sequences: 13736207
Database length: -1841457050
Search time (sec): 2160.120000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000


score_list:
Sequence              Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_est1:AL534526        +  247.00 4562.60  6.8e-245  827  ! AL534526 AL534526 LTI_FL013_FBr
gb_est2:BM461940        +  208.00 3838.14  1.5e-204  913  ! BM461940 AGENCOURT_6418271 NIH_
gb_est1:AW951022        +  147.00 2709.32  1.2e-141  599  ! AW951022 EST363092 MAGE reseque
gb_est1:AU125557        +  147.00 2706.98  1.6e-141  819  ! AU125557 AU125557 NT2RM4 Homo s
gb_est2:BG287359        +  120.00 2204.40  1.5e-113 1007  ! BG287359 602382123F1 NIH_MGC_9
gb_est1:AU127028        +  119.00 2188.04  1.3e-112  751  ! AU127028 AU127028 NT2RP2 Homo s
gb_est2:T10341          +  116.00 2137.09  8.6e-110  399  ! T10341 seq1254 b4HB3MA Cot8-HAF
gb_est1:AL602756        +  115.00 2115.94  1.3e-108  565  ! AL602756 DKFZp686H1516_r1 686 (
gb_est2:BE780308        +  115.00 2112.37  2.1e-108  911  ! BE780308 601468289F1 NIH_MGC_67
gb_est2:BF966141        +  113.00 2078.35  1.6e-106  602  ! BF966141 602286412F1 NIH_MGC_95
gb_est1:AU127381        +  111.00 2038.87  2.6e-104  826  ! AU127381 AU127381 NT2RP2 Homo s
gb_est1:AV649560        +  109.00 2003.85  2.3e-102  624  ! AV649560 AV649560 GLC Homo sapi
gb_est1:AA011446        -  108.00 1987.82  1.8e-101  445  ! AA011446 zi01g12.s1 Soares_feta
gb_est1:AA984697        +  108.00 1987.00  2.0e-101  497  ! AA984697 am89h06.s1 Stratagene
gb_est1:AL602306        +  108.00 1986.62  2.1e-101  523  ! AL602306 DKFZp686J0612_r1 686 (
gb_est1:AU119346        +  107.00 1964.47  3.6e-100  846  ! AU119346 AU119346 HEMBA1 Homo s
gb_est1:AV649613        +   98.00 1799.73  5.3e-91   624  ! AV649613 AV649613 GLC Homo sapi
gb_est1:AU133051        +   94.00 1722.35  1.1e-86   951  ! AU133051 AU133051 NT2RP4 Homo s
gb_est2:BG177093        +   87.00 1592.49  1.9e-79   946  ! BG177093 602313329F1 NIH_MGC_85
gb_est1:AA298723        +   85.00 1563.27  7.9e-78   329  ! AA298723 EST114329 Skin tumor I
gb_est1:AA011445        +   82.00 1505.46  1.3e-74   438  ! AA011445 zi01g12.r1 Soares_feta
gb_est1:AA641754        +   74.00 1360.82  1.5e-66   263  ! AA641754 ns19c06.r1 NCI_CGAP_GC
gb_est2:BM149220        +   64.00 1171.86  5.0e-56   414  ! BM149220 TCAAP2Q11479 Pediatric
gb_est2:BF171567        +   63.00 1157.26  3.3e-55   244  ! BF171567 PCL2964 Myeloma (PCL)
gb_est1:AU127408        +   63.00 1152.21  6.2e-55   479  ! AU127408 AU127408 NT2RP2 Homo s
gb_est2:BM194175        +   62.00 1136.96  4.4e-54   308  ! BM194175 TCAAP1Q11479 Pediatric
gb_est2:BG392166        +   57.00 1034.67  2.2e-48  1098  ! BG392166 602410279F1 NIH_MGC_9
gb_est1:AV648018        +   56.00 1023.91  8.7e-48   387  ! AV648018 AV648018 GLC Homo sapi
gb_est2:N84775          +   53.00  971.36  7.4e-45   255  ! N84775 J0832F Human fetal heart
gb_est1:AA297850        +   53.00  969.73  9.1e-45   317  ! AA297850 EST113482 Jurkat T-cel
gb_est1:AA641875        +   49.00  894.72  1.4e-40   352  ! AA641875 ns19h11.r1 NCI_CGAP_GC
gb_gss:AQ039691         +   49.00  889.70  2.6e-40   689  ! AQ039691 CIT-HSP-2325L17.TV CIT
gb_est2:BG170404        +   48.00  868.63  3.9e-39   965  ! BG170404 602322734F1 NIH_MGC_89
gb_est1:AA243814        +   44.00  800.78  2.3e-35   411  ! AA243814 zr67e07.r1 Soares_NhHM
gb_est1:AA298542        +   42.00  765.80  2.1e-33   309  ! AA298542 EST114158 HSC172 cells
gb_est2:N85243          +   35.00  636.50  3.3e-26   285  ! N85243 J2957F Human fetal heart
gb_est1:AW542198        -   35.00  633.93  4.6e-26   402  ! AW542198 C0154G07-3 NIA Mouse E
gb_est2:BI691090        +   35.00  630.48  7.2e-26   638  ! BI691090 603309866F1 NCI_CGAP_M
gb_est2:BI657919        +   35.00  630.31  7.3e-26   653  ! BI657919 603282147F1 NIH_CG.
gb_est2:BG865023        +   35.00  629.79  7.8e-26   700  ! BG865023 602799491F1 NIH_CG.
gb_est2:BI652204        +   35.00  628.85  8.8e-26   794  ! BI652204 603299029F1 NIH_CG.
gb_est2:BE284569        +   34.00  611.18  8.5e-25   705  ! BE284569 601087184F1 NCI_CG.
gb_est2:BF825247        -   33.00  596.42  5.7e-24   424  ! BF825247 CM1-HN0016-141100-
gb_est2:BI153295        +   30.00  535.99  1.3e-20   802  ! BI153295 602917153F1 NCI_CG.
gb_est2:BI080567        +   30.00  535.31  1.4e-20   878  ! BI080567 602877052F1 NCI_CG.

seq_name: gb_est1:AL534526

seq_documentation_block:
LOCUS       AL534526                827 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL534526 LTI_FL013_FBrn1 Homo sapiens cDNA clone CS0DF004YA06 5
            prime, mRNA sequence.
ACCESSION   AL534526
VERSION     AL534526.1  GI:12798019
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 827)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .827
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DF004YA06"
                     /clone_lib="LTI_FL013_FBrn1"
                     /dev_stage="pooled tissue from post conception fetuses (20
                     week, 24 week and 26 week)"
                     /lab_host="DH10B"
                     /note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand
                     cDNA was primed with a NotI-oligo(dT) primer. Five prime
                     end enriched, double-stranded cDNA was digested with Not I
                     and cloned into the Not I and Eco RV sites of the
                     pCMVSPORT 6 vector.  Library was constructed by Life
                     Technologies.  Contact : Feng Liang Life Technologies, a
                     division of Invitrogen 9800 Medical Center Drive Rockville
                     , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
                     fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      266 a    140 c    192 g    229 t
ORIGIN


alignment_scores:
            Quality:  247.00              Length:    247
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AL534526   ..

 Align seg 1/1  to: AL534526  from: 1  to: 827

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 100

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 150

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 200
```

```
     51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    201 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 250

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    251 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 300

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    301 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 350

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    351 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 400

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    401 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 450

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    451 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 500

    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    501 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 550

    167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    551 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 600

    184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    601 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 650

    201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    651 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 700

    217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    701 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 750

    234 isIleLeuThrArgValAsnArgLysValAlaThrGluPhe 247
        |||||||||||||||||||||||||||||||||||||||
    751 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTT 791

seq_name: gb_est2:BM461940

seq_documentation_block:
LOCUS       BM461940                913 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6418271 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5534002
            5', mRNA sequence.
ACCESSION   BM461940
VERSION     BM461940.1  GI:18510980
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 913)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12219  row: m  column: 11
            High quality sequence stop: 671.
FEATURES             Location/Qualifiers
     source          1. .913
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5534002"
                     /clone_lib="NIH_MGC_71"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
BASE COUNT      285 a    159 c    203 g    265 t      1 others
ORIGIN


alignment_scores:
            Quality:  208.00              Length:    208
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x BM461940   ..

 Align seg 1/1  to: BM461940  from: 1  to: 913

      1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     39 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 88

     17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     89 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 138

     34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    139 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 188

     51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    189 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 238

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    239 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 288

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    289 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 338

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    339 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 388

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    389 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 438

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    439 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 488

    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    489 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 538

    167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    539 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 588
```

```
     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     589 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 638

     201 ProGlyTyrTyrSerTrpArgAsn 208
         |||||||||||||||||||||||
     639 CCTGGTTATTATTCTTGGCGAAAT 662

seq_name: gb_est1:AW951022

seq_documentation_block:
LOCUS       AW951022                599 bp    mRNA    linear   EST 01-JUN-2000
DEFINITION  EST363092 MAGE resequences, MAGA Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW951022
VERSION     AW951022.1  GI:8140687
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 599)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
            ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 22
            Seq primer: Reverse.
FEATURES             Location/Qualifiers
     source          1. .599
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGA"
                     /note="Vector: pBluescriptSKm"
BASE COUNT      218 a     85 c    135 g    160 t      1 others
ORIGIN


alignment_scores:
            Quality:  147.00             Length:    147
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AW951022   ..

 Align seg 1/1  to: AW951022  from: 1  to: 599

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      38 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 87

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      88 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 137

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     138 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 187

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     188 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 237

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||

     238 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 287

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     288 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 337

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     338 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 387

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     388 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 437

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArg 147
         |||||||||||||||||||||||||||||||||||||||
     438 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGT 478

seq_name: gb_est1:AU125557

seq_documentation_block:
LOCUS       AU125557                819 bp    mRNA    linear   EST 23-OCT-2000
DEFINITION  AU125557 NT2RM4 Homo sapiens cDNA clone NT2RM4001791 5', mRNA
            sequence.
ACCESSION   AU125557
VERSION     AU125557.1  GI:10950273
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 819)
  AUTHORS   Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,
            Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,
            Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki
            ,Y., Sugano,S., Isogai,T.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .819
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RM4001791"
                     /clone_lib="NT2RM4"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
                     precursor cells"
BASE COUNT      244 a    149 c    168 g    254 t      4 others
ORIGIN


alignment_scores:
            Quality:  147.00             Length:    147
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AU125557   ..

 Align seg 1/1  to: AU125557  from: 1  to: 819
```

```
    131 AsnGlyProValAspLeuLysLysIleThrAsnPhePheArgGlyAspAr 147
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      3 AATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCG 52

    147 gCysArgSerLeuThrGlyLysProLysLeuPheIleIleGlnAlaCysA 164
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     53 TTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCC 102

    164 rgGlyThrGluLeuAspCysGlyIleGluThrAspSerGlyValAspAsp 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    103 GTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGAT 152

    181 AspMetAlaCysHisLysIleProValGluAlaAspPheLeuTyrAlaTy 197
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    153 GACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATA 202

    197 rSerThrAlaProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerT 214
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    203 CTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCT 252

    214 rpPheIleGlnSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeu 230
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    253 GGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTT 302

    231 GluPheMetHisIleLeuThrArgValAsnArgLysValAlaThrGluPh 247
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    303 GAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATT 352

    247 eGluSerPheSerPheAspAlaThrPheHisAlaLysLysGlnIleProC 264
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    353 TGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCAT 402

    264 ysIleValSerMetLeuThrLysGluLeuTyrPheTyrHis 277
        |||||||||||||||||||||||||||||||||||||
    403 GTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 443

seq_name: gb_est2:BG287359

seq_documentation_block:
LOCUS       BG287359                1007 bp    mRNA    linear   EST 21-FEB-2001
DEFINITION  602382123F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4499885 5',
            mRNA sequence.
ACCESSION   BG287359
VERSION     BG287359.1  GI:13041121
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1007)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10364  row: m  column: 06
            High quality sequence stop: 655.
FEATURES             Location/Qualifiers
     source          1. .1007
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4499885"
                     /clone_lib="NIH_MGC_93"
                     /tissue_type="transitional cell papilloma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.7 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      265 a    213 c    235 g    294 t
ORIGIN


alignment_scores:
           Quality:  120.00              Length:    120
             Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x BG287359   ..

 Align seg 1/1  to: BG287359  from: 1  to: 1007

    158 PheIleIleGlnAlaCysArgGlyThrGluLeuAspCysGlyIleGluTh 174
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      1 TTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGAC 50

    174 rAspSerGlyValAspAspAspMetAlaCysHisLysIleProValGluA 191
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 AGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGG 100

    191 laAspPheLeuTyrAlaTyrSerThrAlaProGlyTyrTyrSerTrpArg 207
        |||||||||||||||||||||||||||||||||||||||||||||||||
    101 CCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGA 150

    208 AsnSerLysAspGlySerTrpPheIleGlnSerLeuCysAlaMetLeuLy 224
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    151 AATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAA 200

    224 sGlnTyrAlaAspLysLeuGluPheMetHisIleLeuThrArgValAsnA 241
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    201 ACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACC 250

    241 rgLysValAlaThrGluPheGluSerPheSerPheAspAlaThrPheHis 257
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    251 GAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCAT 300

    258 AlaLysLysGlnIleProCysIleValSerMetLeuThrLysGluLeuTy 274
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    301 GCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTA 350

    274 rPheTyrHis 277
        |||||||||
    351 TTTTTATCAC 360

seq_name: gb_est1:AU127028

seq_documentation_block:
LOCUS       AU127028                751 bp    mRNA    linear   EST 23-OCT-2000
DEFINITION  AU127028 NT2RP2 Homo sapiens cDNA clone NT2RP2000417 5', mRNA
            sequence.
ACCESSION   AU127028
VERSION     AU127028.1  GI:10951744
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 751)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
```

```
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES              Location/Qualifiers
     source           1. .751
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="NT2RP2000417"
                      /clone_lib="NT2RP2"
                      /cell_type="teratocarcinoma"
                      /cell_line="NT2"
                      /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                      cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT      242 a    129 c    175 g    194 t     11 others
ORIGIN


alignment_scores:
            Quality:  119.00             Length:    119
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AU127028   ..

 Align seg 1/1  to: AU127028  from: 1  to: 751

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
      97 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 146

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     147 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 196

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     197 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 246

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     247 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 296

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     297 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 346

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     347 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 396

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     397 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 446

     117 lLeuLeu 119
         |||||||
     447 GCTTCTG 453

seq_name: gb_est2:T10341

seq_documentation_block:
LOCUS       T10341                  399 bp    mRNA    linear   EST 07-JUN-1994
DEFINITION  seq1254 b4HB3MA Cot8-HAP-Ft Homo sapiens cDNA clone
            b4HB3MA-COT8-HAP-Ft280 5' similar to similar to Human Cysteine
            proteinase and to Mouse interleukin 1-beta convertase, mRNA
            sequence.
ACCESSION   T10341
VERSION     T10341.1  GI:471690
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 399)
  AUTHORS   Soares,M.B., Bonaldo,M.F., Jelenc,P., Su,L., Lawton,L and
            Efstratiadis,A.
  TITLE     Construction and characterization of a normalized cDNA library
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 91, 9228-9232 (1994)
  MEDLINE   95023884
COMMENT     Contact: Bento Soares
            Columbia University
            Department of Psychiatry,722 W 168th Street, Unit #41, New York,NY
            10032
            Tel: 2129602313
            Fax: 2127813577
            Email: cuc@cuccfa.ccc.columbia.edu
            sense, Ampicillin
            Seq primer: M13 Reverse.
FEATURES              Location/Qualifiers
     source           1. .399
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="b4HB3MA-COT8-HAP-Ft280"
                      /clone_lib="b4HB3MA Cot8-HAP-Ft"
                      /lab_host="E. coli"
                      /note="Vector: Lafmid BA; Site_1: Not I; Site_2: Hind III;
                      Size-selected cDNA from polyA+ RNA from human brain.
                      3-month old neonate.This library is the result of an
                      attempted normalization of library b4HB3MA. "
BASE COUNT      111 a     84 c     94 g    110 t
ORIGIN


alignment_scores:
            Quality:  116.00             Length:    116
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x T10341   ..

 Align seg 1/1  to: T10341  from: 1  to: 399

     137 LysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThrGl 153
         |||||||||||||||||||||||||||||||||||||||||||||||||||
      20 AAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGG 69

     153 yLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAspC 170
         |||||||||||||||||||||||||||||||||||||||||||||||||||
      70 AAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACT 119

     170 ysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisLys 186
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     120 GTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 169

     187 IleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGlyTy 203
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     170 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTA 219

     203 rTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLeuC 220
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     220 TTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTT 269

     220 ysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleLeu 236
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     270 GTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTT 319

     237 ThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
         |||||||||||||||||||||||||||||||||||||||||||||||||
     320 ACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 367
```

```
seq_name: gb_est1:AL602756

seq_documentation_block:
LOCUS       AL602756                565 bp    mRNA    linear   EST 14-AUG-2001
DEFINITION  DKFZp686H1516_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686H1516 5', mRNA sequence.
ACCESSION   AL602756
VERSION     AL602756.1  GI:15166262
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 565)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,W., Weil,B. and Wiemann,S.
  TITLE     EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B. and Wiemann,S.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Wambutt R
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686H1516) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .565
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686H1516"
                     /clone_lib="686 (synonym: hlcc3)"
                     /tissue_type="human skeletal muscle"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
BASE COUNT      174 a    102 c    136 g    153 t
ORIGIN


alignment_scores:
            Quality:  115.00             Length:    115
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AL602756   ..

 Align seg 1/1  to: AL602756  from: 1  to: 565

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     219 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 268

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     269 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 318

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     319 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 368

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     369 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 418

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||

     419 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 468

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     469 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 518

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheVal 115
         |||||||||||||||||||||||||||||||||||||||||||||
     519 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 563

seq_name: gb_est2:BE780308

seq_documentation_block:
LOCUS       BE780308                911 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601468289F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3871390 5',
            mRNA sequence.
ACCESSION   BE780308
VERSION     BE780308.1  GI:10201506
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 911)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9624  row: a  column: 23
            High quality sequence stop: 768.
FEATURES             Location/Qualifiers
     source          1. .911
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3871390"
                     /clone_lib="NIH_MGC_67"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      287 a    175 c    205 g    244 t
ORIGIN


alignment_scores:
            Quality:  115.00             Length:    189
              Ratio:  0.608                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x BE780308   ..

 Align seg 1/1  to: BE780308  from: 1  to: 911

       5 GluAsnSerValAspSerLysSerIleLysAsnLeuGluProLysIleIl 21
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      14 GAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCAT 63

      21 eHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnSerTyrL 38
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      64 ACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATA 113

      38 ysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsnLysAsn 54
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 114 AAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAAT 163

  55 PheHisLysSerThrGlyMetThrSerArgSerGlyThrAspValAspAl 71
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 164 TTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGC 213

  71 aAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValArgAsnL 88
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 214 AGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATA 263

  88 ysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAspValSer 104
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 264 AAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCT 313

 105 LysGluAspHisSerLysArgSerSerPheValCysValLeuLeuSerHi 121
     ||||||||||||||||||||||    ||||||||||||||||||||||||
 314 AAAGAAGATCACAGCAAAAGGAGCAG.TTTGTTTGTGTGCTTCTGAGCCA 362

 121 sGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspLeuLysL 138
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 363 TGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAA 412

 138 ysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThrGlyLys 154
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 413 AAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAA 462

 155 ProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAspCysGl 171
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 463 CCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGG 512

 171 yIleGluThrAspSerGlyValAspAspAspMetAlaCysHisLysIleP 188
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 513 CATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATAC 562

 188 roValGluAlaAspPhe 193
     |||||||||||||||||
 563 CAGTGGAGGCCGACTTC 579

seq_name: gb_est2:BF966141

seq_documentation_block:
LOCUS       BF966141               602 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  602286412F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4375323 5',
            mRNA sequence.
ACCESSION   BF966141
VERSION     BF966141.1  GI:12333356
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 602)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10040  row: g  column: 04
            High quality sequence stop: 599.
FEATURES             Location/Qualifiers
     source          1. .602
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4375323"
                     /clone_lib="NIH_MGC_95"
                     /tissue_type="hippocampus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.5 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for  full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health).  Note: this is a NIH_MGC Library."
BASE COUNT      206 a     97 c    149 g    150 t
ORIGIN


alignment_scores:
            Quality:  113.00              Length:    113
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x BF966141   ..

 Align seg 1/1  to: BF966141  from: 1  to: 602

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 150

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 200

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 250

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 300

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 350

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 400

     101 ArgAspValSerLysGluAspHisSerLysArgSerSer 113
         |||||||||||||||||||||||||||||||||||||
     401 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGT 439

seq_name: gb_est1:AU127381

seq_documentation_block:
LOCUS       AU127381               826 bp    mRNA    linear   EST 23-OCT-2000
DEFINITION  AU127381 NT2RP2 Homo sapiens cDNA clone NT2RP2001169 5', mRNA
            sequence.
ACCESSION   AU127381
VERSION     AU127381.1  GI:10952097
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 826)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
```

```
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .826
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP2001169"
                     /clone_lib="NT2RP2"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                     cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT      254 a    152 c    188 g    226 t      6 others
ORIGIN


alignment_scores:
           Quality:  111.00            Length:    111
             Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AU127381   ..

 Align seg 1/1  to: AU127381  from: 1  to: 826

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     233 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 282

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     283 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 332

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     333 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 382

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     383 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 432

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     433 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 482

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     483 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 532

     101 ArgAspValSerLysGluAspHisSerLysArg 111
         |||||||||||||||||||||||||||||||
     533 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGG 565

seq_name: gb_est1:AV649560

seq_documentation_block:
LOCUS       AV649560                624 bp    mRNA    linear   EST 15-JAN-2002
DEFINITION  AV649560 GLC Homo sapiens cDNA clone GLCBVC04 3', mRNA sequence.
ACCESSION   AV649560
VERSION     AV649560.1  GI:9870574
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
            Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
            Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
            Hu,G., Gu,J., Chen,Z. and Han,Z.
  TITLE     Insight into hepatocellular carcinogenesis at transcriptome level
            by comparing gene expression profiles of hepatocellular carcinoma
            with those of corresponding noncancerous liver
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
  MEDLINE   21625106
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="GLCBVC04"
                     /clone_lib="GLC"
                     /tissue_type="corresponding non cancerous liver tissue"
                     /dev_stage="Adult"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      182 a    116 c    156 g    166 t      4 others
ORIGIN


alignment_scores:
           Quality:  109.00            Length:    109
             Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AV649560   ..

 Align seg 1/1  to: AV649560  from: 1  to: 624

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     216 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 265

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     266 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 315

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     316 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 365

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     366 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 415

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     416 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 465

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     466 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 515

     101 ArgAspValSerLysGluAspHisSer 109
         |||||||||||||||||||||||||||
     516 CGTGATGTTTCTAAAGAAGATCACAGC 542
```

```
seq_name: gb_est1:AA011446

seq_documentation_block:
LOCUS       AA011446                 445 bp    mRNA    linear   EST 29-JUL-1996
DEFINITION  zi01g12.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA
            clone IMAGE:429574 3' similar to SW:CP32_HUMAN P42574 APOPAIN
            PRECURSOR ;, mRNA sequence.
ACCESSION   AA011446
VERSION     AA011446.1  GI:1472543
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 445)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: mob.REGA+ET
            High quality sequence stop: 395.
FEATURES             Location/Qualifiers
     source          1. .445
                     /organism="Homo sapiens"
                     /db_xref="GDB:1329343"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:429574"
                     /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     This is a subtracted version of the original Soares fetal
                     liver spleen 1NFLS library.  1st strand cDNA was primed
                     with a Pac I - oligo(dT) primer [5'
                     AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      127 a     98 c     90 g    128 t      2 others
ORIGIN


alignment_scores:
            Quality:  108.00             Length:    108
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AA011446/rev   ..

 Align seg 1/1  to reverse of: AA011446  from: 1  to: 445

     162 AlaCysArgGlyThrGluLeuAspCysGlyIleGluThrAspSerGlyVa 178
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     326 GCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGT 277

     178 lAspAspAspMetAlaCysHisLysIleProValGluAlaAspPheLeuT 195
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     276 TGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGT 227

     195 yrAlaTyrSerThrAlaProGlyTyrTyrSerTrpArgAsnSerLysAsp 211
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     226 ATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGAT 177

     212 GlySerTrpPheIleGlnSerLeuCysAlaMetLeuLysGlnTyrAlaAs 228
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     176 GGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGA 127

     228 pLysLeuGluPheMetHisIleLeuThrArgValAsnArgLysValAlaT 245
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     126 CAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAA 77

     245 hrGluPheGluSerPheSerPheAspAlaThrPheHisAlaLysLysGln 261
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      76 CAGAATTTGAGTCCTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAG 27

     262 IleProCysIleValSerMetLeu 269
         |||||||||||||||||||||||||
      26 ATTCCATGTATTGTTTCCATGCTC 3

seq_name: gb_est1:AA984697

seq_documentation_block:
LOCUS       AA984697                 497 bp    mRNA    linear   EST 27-MAY-1998
DEFINITION  am89h06.s1 Stratagene schizo brain S11 Homo sapiens cDNA clone
            IMAGE:1630331 3' similar to SW:ICE3_HUMAN P42574 APOPAIN PRECURSOR
            ;, mRNA sequence.
ACCESSION   AA984697
VERSION     AA984697.1  GI:3163222
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 497)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 372.
FEATURES             Location/Qualifiers
     source          1. .497
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1630331"
                     /clone_lib="Stratagene schizo brain S11"
                     /sex="male"
                     /tissue_type="schizophrenic brain S-11 frontal lobe"
                     /dev_stage="34 years old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Vector: Bluescript SK-; Site_1: EcoRI;  Library
                     constructed from S-11 frontal lobe, male, 34 years old,
                     50% caucasian, 50% Aleutian.  Schizophrenic suicide.
                     Random primed into EcoRI site of ZAP II Vector.  Mass
                     excised. Custom library. Avg insert length 1.4kb.
                     Material obtained by Johnston N., Torrey, E.F., Yolken R.,
                     and the Stanley Neuropathology Consortium - Analysis of
                     RNAs from the Brains of Individuals with Psychiatric
                     Diseases (Unpublished) Stanley Neurovirology Laboratory,
                     Johns Hopkins School of Medicine, Baltimore MD."
```

```
BASE COUNT      153 a    101 c    113 g    130 t
ORIGIN


alignment_scores:
           Quality:  108.00             Length:    108
             Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AA984697   ..

 Align seg 1/1  to: AA984697  from: 1  to: 497

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     174 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 223

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     224 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 273

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     274 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 323

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     324 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 373

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     374 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 423

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     424 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 473

     101 ArgAspValSerLysGluAspHis 108
         |||||||||||||||||||||||||
     474 CGTGATGTTTCTAAAGAAGATCAC 497

seq_name: gb_est1:AL602306

seq_documentation_block:
LOCUS       AL602306               523 bp    mRNA    linear   EST 14-AUG-2001
DEFINITION  DKFZp686J0612_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686J0612 5', mRNA sequence.
ACCESSION   AL602306
VERSION     AL602306.1  GI:15165812
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 523)
  AUTHORS   Poustka,A., Wellenreuther,R., Mewes,H.W., Weil,B. and Wiemann,S.).
  TITLE     EST (Poustka,A., Wellenreuther,R., Mewes,H.W., Weil,B. and Wiemann
            ,S.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Poustka A.J.
            Department Lehrach
            Max-Planck-Institute for Molecular Genetics
            Ihnestrasse 73, 14195 Berlin, Germany
            Tel: +49-30-84131623
            Fax: +49-30-84131128
            Email: poustka@mpimg-berlin-dahlem.mpg.de
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686J0612) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES              Location/Qualifiers
     source           1. .523
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="DKFZp686J0612"
                      /clone_lib="686 (synonym: hlcc3)"
                      /tissue_type="human skeletal muscle"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                      cDNA-collection"
BASE COUNT      161 a     95 c    132 g    135 t
ORIGIN


alignment_scores:
           Quality:  108.00             Length:    108
             Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AL602306   ..

 Align seg 1/1  to: AL602306  from: 1  to: 523

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     199 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 248

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     249 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 298

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     299 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 348

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     349 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 398

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     399 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 448

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     449 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 498

     101 ArgAspValSerLysGluAspHis 108
         |||||||||||||||||||||||||
     499 CGTGATGTTTCTAAAGAAGATCAC 522
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2002, 23:58:34 ; Search time 3312.14 Seconds
(without alignments)
8662.159 Million cell updates/sec

Title: US-09-613-508-1
Perfect score: 1371
Sequence: 1 GCACGAGAAACTTTGCTGTG.....TAATTTCAGGAAAATGTATT 1371

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1371	100.0	1371	6	AR103058	AR103058 Sequence
2	849	61.9	2309	6	AX134417	AX134417 Sequence
3	849	61.9	2309	9	HSU37448	U37448 Human Mch3
4	848	61.9	2361	9	BC015799	BC015799 Homo sapi
5	833	60.8	2335	9	HSU40281	U40281 Human cyste
6	798	58.2	2314	9	HSU67320	U67320 Human Lice2
7	798	58.2	2602	9	HSU67319	U67319 Human Lice2
8	798	58.2	2661	9	HSU67206	U67206 Homo sapien
9	717	52.3	912	9	HSU39613	U39613 Human cyste
10	433	31.6	2378	9	HSU37449	U37449 Human Mch3
11	246	17.9	155904	2	AL627395	AL627395 Homo sapi
c 12	246	17.9	174776	9	AL592546	AL592546 Human DNA
13	32	2.3	964	10	D86353	D86353 Mouse mRNA
14	32	2.3	1063	10	AF072124	AF072124 Rattus no
15	32	2.3	2006	10	MMU67321	U67321 Mus musculu
16	32	2.3	2305	10	BC005428	BC005428 Mus muscu
17	32	2.3	2372	10	MMCASP7	Y13088 M.musculus
c 18	28	2.0	28	6	AR091202	AR091202 Sequence
19	26	1.9	26	6	AR091201	AR091201 Sequence
20	26	1.9	1939	10	MAU47332	U47332 Mesocricetu
21	24	1.8	490	9	HSAY18953	Y18953 Homo sapien
22	23	1.7	69227	2	AC083985	AC083985 Homo sapi
c 23	23	1.7	193973	9	AC097499	AC097499 Homo sapi
24	22	1.6	22	6	AR103064	AR103064 Sequence
c 25	22	1.6	53	6	AR103065	AR103065 Sequence
26	21	1.5	31	6	AR103060	AR103060 Sequence
c 27	21	1.5	31	6	AR103061	AR103061 Sequence
28	21	1.5	1659	9	HSCD89S1	AF091543 Homo sapi
c 29	21	1.5	11862	8	ATU53502	U53502 Arabidopsis
30	21	1.5	31226	3	U39998	U39998 Caenorhabdi
c 31	21	1.5	89811	8	AC018849	AC018849 Arabidops
c 32	21	1.5	119405	2	AL669909	AL669909 Mus muscu
33	21	1.5	132990	8	AC018848	AC018848 Arabidops
c 34	21	1.5	175878	2	AC097080	AC097080 Rattus no
35	21	1.5	177160	9	AC019089	AC019089 Homo sapi
36	21	1.5	184659	2	AC021534	AC021534 Homo sapi
37	21	1.5	193260	2	AC025598	AC025598 Homo sapi
38	21	1.5	209515	2	AC110006	AC110006 Homo sapi
39	21	1.5	214702	9	AC020689	AC020689 Homo sapi
40	20	1.5	1096	3	AF130469	AF130469 Drosophil
41	20	1.5	1153	10	S53376	S53376 SC1 [mice,
c 42	20	1.5	1362	9	AK000559	AK000559 Homo sapi
c 43	20	1.5	1440	9	BC001442	BC001442 Homo sapi
44	20	1.5	1611	10	BC004617	BC004617 Mus muscu
c 45	20	1.5	1628	6	AX274884	AX274884 Sequence

ALIGNMENTS

RESULT 1
AR103058
LOCUS AR103058 1371 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6087150.
ACCESSION AR103058
VERSION AR103058.1 GI:12814646
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1371)
AUTHORS He,W.W., Rosen,C.A., Hudson,P.L. and Hastings,G.A.
TITLE Interleukin-1 .beta. converting enzyme like apoptosis protease-3 and 4
JOURNAL Patent: US 6087150-A 1 11-JUL-2000;
FEATURES Location/Qualifiers
source 1. .1371
/organism="unknown"
BASE COUNT 355 a 357 c 354 g 305 t
ORIGIN

Query Match 100.0%; Score 1371; DB 6; Length 1371;

```
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1371;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCACGAGAAACTTTGCTGTGCGCGTTCTCCCGCGCGCGGGCTCAACTTTGTAGAGCGAGG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GCACGAGAAACTTTGCTGTGCGCGTTCTCCCGCGCGCGGGCTCAACTTTGTAGAGCGAGG 60

Qy      61 GGCCAACTTGGCAGAGCGCGCGGCCAGCTTTGCAGAGAGCGCCCTCCAGGGACTATGCGT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GGCCAACTTGGCAGAGCGCGCGGCCAGCTTTGCAGAGAGCGCCCTCCAGGGACTATGCGT 120

Qy     121 GCGGGGACACGGGTCGCTTTGGGCTCTTCCACCCCTGCGGAGCGCACTACCCCGAGCCAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCGGGGACACGGGTCGCTTTGGGCTCTTCCACCCCTGCGGAGCGCACTACCCCGAGCCAG 180

Qy     181 GGGCGGTGCAAGCCCCGCCCGGCCCTACCCAGGGCGGCTCCTCCCTCCGCAGCGCCGAGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GGGCGGTGCAAGCCCCGCCCGGCCCTACCCAGGGCGGCTCCTCCCTCCGCAGCGCCGAGA 240

Qy     241 CTTTTAGTTTCGCTTTCGCTAAAGGGGCCCCAGACCCTTGCTGCGGAGCGACGGAGAGAG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTTTTAGTTTCGCTTTCGCTAAAGGGGCCCCAGACCCTTGCTGCGGAGCGACGGAGAGAG 300

Qy     301 ACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGGGCTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGGGCTGT 360

Qy     361 ATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGAC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGAC 420

Qy     421 CGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCC 480

Qy     481 ATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTG 540

Qy     541 GGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGA 600

Qy     601 AACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGAC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGAC 660

Qy     661 GTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCT 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCT 720

Qy     721 GAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAA 780

Qy     781 AATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 AATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGG 840

Qy     841 GGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGA 900

Qy     901 GGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCT 960

Qy     961 AATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCA 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 AATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCA 1020

Qy    1021 GGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1021 GGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCC 1080

Qy    1081 ATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 ATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGA 1140

Qy    1141 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCA 1200

Qy    1201 GATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCA 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 GATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCA 1260

Qy    1261 GGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGC 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGC 1320

Qy    1321 TCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 TCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371


RESULT   2
AX134417
LOCUS       AX134417               2309 bp    DNA     linear   PAT 29-MAY-2001
DEFINITION  Sequence 13 from Patent WO0129232.
ACCESSION   AX134417
VERSION     AX134417.1  GI:14271063
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2309)
  AUTHORS   Cordell,B. and Li,Y.
  TITLE     Functional cloning of genes encoding proteins/enzymes involved in
            proteolytic cleavage
  JOURNAL   Patent: WO 0129232-A 13 26-APR-2001;
            Scios Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2309
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             44. .955
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC39525.1"
                     /db_xref="GI:14271064"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      680 a    453 c    531 g    645 t
ORIGIN


  Query Match           61.9%;  Score 849;  DB 6;  Length 2309;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1019;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy     351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 114

Qy     411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 174

Qy     471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     175 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 234

Qy     531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     235 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 294

Qy     591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     295 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 354

Qy     651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     355 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 414

Qy     711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     415 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 474

Qy     771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     475 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 534

Qy     831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     535 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 594

Qy     891 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 950
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     595 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 654

Qy     951 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 1010
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     655 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 714

Qy    1011 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 1070
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     715 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 774

Qy    1071 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCA 1130
           |||||||||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db     775 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCA 833

Qy    1131 GGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 1190
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     834 GGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 893

Qy    1191 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAAT 1250
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     894 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAAT 953

Qy    1251 AGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACAT 1310
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     954 AGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACAT 1013

Qy    1311 TTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTAT 1370
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1014 TTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTAT 1073

Qy    1371 T 1371
           |
Db    1074 T 1074
```

```
RESULT   3
HSU37448
LOCUS       HSU37448                2309 bp    mRNA    linear   PRI 13-DEC-1995
DEFINITION  Human Mch3 isoform alpha (Mch3) mRNA, complete cds.
ACCESSION   U37448
VERSION     U37448.1  GI:1117846
KEYWORDS    Mch3 isoform alpha.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2309)
  AUTHORS   Fernandes-Alnemri,T., Litwack,G. and Alnemri,E.S.
  TITLE     CPP32, a novel human apoptotic protein with homology to
            Caenorhabditis elegans cell death protein Ced-3 and mammalian
            interleukin-1 beta-converting enzyme
  JOURNAL   J. Biol. Chem. 269 (49), 30761-30764 (1994)
  MEDLINE   95074098
REFERENCE   2  (bases 1 to 2309)
  AUTHORS   Fernandes-Alnemri,T., Takahashi,A., Armstrong,R., Krebs,J.,
            Fritz,L., Tomaselli,K.J., Wang,L., Yu,Z., Croce,C.M., Salveson,G.,
            Earnshaw,W.C., Litwack,G. and Alnemri,E.S.
  TITLE     Mch3, a novel human apoptotic cysteine protease highly related to
            CPP32
  JOURNAL   Cancer Res. 55 (24), 6045-6052 (1995)
  MEDLINE   96105019
REFERENCE   3  (bases 1 to 2309)
  AUTHORS   Alnemri,E.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-1995) Emad S. Alnemri, Pharmacology, Thomas
            Jefferson University, Jefferson Cancer Institute, 233, S. Tenth
            Street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .2309
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="T50-8A.1"
                     /cell_line="Jurkat"
                     /cell_type="T-lymphocyte"
     gene            1. .2309
                     /gene="Mch3"
     CDS             44. .955
                     /gene="Mch3"
                     /function="cysteine protease"
                     /note="new member of ICE-like cysteine protease family"
                     /codon_start=1
                     /evidence=experimental
                     /product="Mch3 isoform alpha"
                     /protein_id="AAC50303.1"
                     /db_xref="GI:1117847"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
     polyA_site      2309
                     /gene="Mch3"
BASE COUNT      680 a    453 c    531 g    645 t
ORIGIN
```

Query Match 61.9%; Score 849; DB 9; Length 2309;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1019; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
Qy     351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 114

Qy     411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 174

Qy     471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     175 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 234

Qy     531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     235 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 294

Qy     591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
```

```
Db     295 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 354

Qy     651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     355 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 414

Qy     711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     415 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 474

Qy     771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     475 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 534

Qy     831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     535 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 594

Qy     891 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 950
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     595 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 654

Qy     951 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 1010
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     655 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 714

Qy    1011 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 1070
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     715 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 774

Qy    1071 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCA 1130
           |||||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db     775 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCA 833

Qy    1131 GGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 1190
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     834 GGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 893

Qy    1191 AGAAGAAGCAGATCCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAAT 1250
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     894 AGAAGAAGCAGATCCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAAT 953

Qy    1251 AGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACAT 1310
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     954 AGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACAT 1013

Qy    1311 TTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTAT 1370
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1014 TTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTAT 1073

Qy    1371 T 1371
           |
Db    1074 T 1074
```

```
RESULT   4
BC015799
LOCUS       BC015799               2361 bp    mRNA    linear   PRI 29-OCT-2001
DEFINITION  Homo sapiens, caspase 7, apoptosis-related cysteine protease, clone
            MGC:9176 IMAGE:3921455, mRNA, complete cds.
ACCESSION   BC015799
VERSION     BC015799.1  GI:16041818
KEYWORDS    MGC.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2361)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-OCT-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 15 Row: n Column: 12
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 15718696.
FEATURES              Location/Qualifiers
     source           1. .2361
                      /organism="Homo sapiens"
                      /db_xref="LocusID:840"
                      /db_xref="taxon:9606"
                      /clone="MGC:9176 IMAGE:3921455"
                      /tissue_type="Skin, melanotic melanoma."
                      /clone_lib="NIH_MGC_72"
                      /lab_host="DH10B"
                      /note="Vector: pCMV-SPORT6"
     CDS              79. .990
                      /codon_start=1
                      /product="caspase 7, apoptosis-related cysteine protease"
                      /protein_id="AAH15799.1"
                      /db_xref="GI:16041819"
                      /translation="MADEQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                      MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                      SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                      DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                      FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                      SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      704 a    464 c    545 g    648 t
ORIGIN
```

```
  Query Match           61.9%;  Score 848;  DB 9;  Length 2361;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 1018;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy     352 CAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCT 411
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      91 CAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCT 150

Qy     412 AAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACC 471
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151 AAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACC 210

Qy     472 ATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTT 531
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     211 ATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTT 270

Qy     532 GAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATG 591
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271 GAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATG 330

Qy     592 GGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTG 651
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     331 GGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTG 390
```

```
Qy    652 GGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAA 711
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    391 GGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAA 450

Qy    712 AAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCAT 771
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    451 AAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCAT 510

Qy    772 GGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCC 831
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    511 GGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCC 570

Qy    832 CACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAG 891
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    571 CACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAG 630

Qy    892 GCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGAC 951
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    631 GCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGAC 690

Qy    952 ACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCC 1011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    691 ACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCC 750

Qy   1012 ACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCC 1071
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    751 ACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCC 810

Qy   1072 CTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAG 1131
          |||||||||||||||||||||||||||||||||||||||||||||| ||||| ||||||
Db    811 CTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCAG 869

Qy   1132 GGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGA 1191
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    870 GGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGA 929

Qy   1192 GAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATA 1251
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    930 GAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATA 989

Qy   1252 GCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATT 1311
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    990 GCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATT 1049

Qy   1312 TTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1050 TTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1109
```

RESULT 5
HSU40281
LOCUS HSU40281 2335 bp mRNA linear PRI 27-JAN-1996
DEFINITION Human cysteine protease CMH-1 mRNA, complete cds.
ACCESSION U40281
VERSION U40281.1 GI:1167554
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2335)
AUTHORS Lippke,J.A., Gu,Y., Sarnecki,C., Caron,P.R. and Su,M.S.
TITLE Identification and characterization of CPP32/Mch2 homolog 1, a novel cysteine protease similar to CPP32
JOURNAL J. Biol. Chem. 271 (4), 1825-1828 (1996)
MEDLINE 96147144
REFERENCE 2 (bases 1 to 2335)
AUTHORS Caron,P.R.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-1995) Paul R. Caron, Vertex Pharmaceuticals, 40 Allston St., Cambridge, MA 02139, USA
FEATURES Location/Qualifiers

```
     source          1. .2335
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="IMAGE consortium clone 72778"
     CDS             48. .959
                     /function="cysteine protease"
                     /note="similar to CPP32 and to MCH2; related to Ced3 and
                     interleukin-1 beta converting enzyme"
                     /codon_start=1
                     /product="CMH-1"
                     /protein_id="AAC50352.1"
                     /db_xref="GI:1167555"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT 700 a 454 c 535 g 646 t
ORIGIN
```

```
Query Match 60.8%; Score 833; DB 9; Length 2335;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1003; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy    367 GAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCC 426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 GAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCC 134

Qy    427 TCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAG 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 TCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAG 194

Qy    487 ACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAA 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 ACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAA 254

Qy    547 TGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGA 606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 TGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGA 314

Qy    607 ACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATT 666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 ACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATT 374

Qy    667 GTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAG 726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 GTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAG 434

Qy    727 GACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTA 786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    435 GACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTA 494

Qy    787 ATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGAT 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    495 ATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGAT 554

Qy    847 AGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACC 906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    555 AGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACC 614

Qy    907 GAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCT 966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    615 GAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCT 674

Qy    967 CGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTAT 1026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    675 CGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTAT 734

Qy   1027 TACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTG 1086
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    735 TACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTG 794
```

```
Qy   1087 GAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGACAGAGT 1146
          ||||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||
Db    795 GAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCAGGGTGAATGACAGAGT 853

Qy   1147 TGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGATCCC 1206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    854 TGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGATCCC 913

Qy   1207 CTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCAGGGGTA 1266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    914 CTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCAGGGGTA 973

Qy   1267 CATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGCTCTTGA 1326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    974 CATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGCTCTTGA 1033

Qy   1327 AATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1034 AATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1078
```

```
RESULT   6
HSU67320
LOCUS       HSU67320                2314 bp    mRNA    linear   PRI 19-MAR-1997
DEFINITION  Human Lice2 gamma cysteine protease mRNA, complete cds.
ACCESSION   U67320
VERSION     U67320.1  GI:1894914
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2314)
  AUTHORS   Juan,T.S., McNiece,I.K., Argento,J.M., Jenkins,N.A., Gilbert,D.J.,
            Copeland,N.G. and Fletcher,F.A.
  TITLE     Identification and mapping of Casp7, a cysteine protease resembling
            CPP32 beta, interleukin-1 beta converting enzyme, and CED-3
  JOURNAL   Genomics 40 (1), 86-93 (1997)
  MEDLINE   97224489
REFERENCE   2  (bases 1 to 2314)
  AUTHORS   Juan,T.S.C., McNiece,I.K., Argento,J.M., Jenkins,N.A.,
            Gilbert,D.J., Copeland,N.G. and Fletcher,F.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-1996) Developmental Hematology, Amgen, Inc., 1840
            DeHavilland Dr., Thousand Oaks, CA 91320, USA
FEATURES             Location/Qualifiers
     source          1. .2314
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             40. .951
                     /note="similar to ICE and CED-3 cysteine protease"
                     /codon_start=1
                     /product="Lice2 gamma cysteine protease"
                     /protein_id="AAC51153.1"
                     /db_xref="GI:1894915"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      680 a    459 c    526 g    648 t      1 others
ORIGIN
```

```
  Query Match           58.2%;  Score 798;  DB 9;  Length 2314;
  Best Local Similarity  99.7%;  Pred. No. 0;
  Matches 1018;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy    351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     51 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 110

Qy    411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    111 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 170

Qy    471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    171 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 230

Qy    531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    231 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 290

Qy    591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    291 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 350

Qy    651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    351 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 410

Qy    711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    411 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 470

Qy    771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    471 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 530

Qy    831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    531 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 590

Qy    891 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 950
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    591 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 650

Qy    951 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 1010
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    651 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 710

Qy   1011 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 1070
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    711 CACGGTTCCAGGCTATTACTCATGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 770

Qy   1071 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCA 1130
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db    771 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCA 829

Qy   1131 GGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 1190
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    830 GGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 889

Qy   1191 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAAT 1250
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    890 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAAT 949

Qy   1251 AGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACAT 1310
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    950 AGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACAT 1009

Qy   1311 TTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTAT 1370
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1010 TTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTAT 1069

Qy   1371 T 1371
          |
Db   1070 T 1070
```

```
RESULT   7
HSU67319
LOCUS       HSU67319                2602 bp    mRNA    linear   PRI 19-MAR-1997
```

```
DEFINITION  Human Lice2 beta cysteine protease mRNA, complete cds.
ACCESSION   U67319
VERSION     U67319.1  GI:1894912
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2602)
  AUTHORS   Juan,T.S., McNiece,I.K., Argento,J.M., Jenkins,N.A., Gilbert,D.J.,
            Copeland,N.G. and Fletcher,F.A.
  TITLE     Identification and mapping of Casp7, a cysteine protease resembling
            CPP32 beta, interleukin-1 beta converting enzyme, and CED-3
  JOURNAL   Genomics 40 (1), 86-93 (1997)
  MEDLINE   97224489
REFERENCE   2  (bases 1 to 2602)
  AUTHORS   Juan,T.S.C., McNiece,I.K., Argento,J.M., Jenkins,N.A.,
            Gilbert,D.J., Copeland,N.G. and Fletcher,F.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-1996) Developmental Hematology, Amgen, Inc., 1840
            DeHavilland Dr., Thousand Oaks, CA 91320, USA
FEATURES             Location/Qualifiers
     source          1. .2602
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             229. .1239
                     /note="similar to ICE and CED-3 cysteine protease"
                     /codon_start=1
                     /product="Lice2 beta cysteine protease"
                     /protein_id="AAC51152.1"
                     /db_xref="GI:1894913"
                     /translation="MDCVGWPPGRKWHLEKNTSCGGSSGICASYVTQMADDQGCIEEQ
                     GVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVTMRSIKTTRDRVPTYQYNMNFEKLGK
                     CIIINNKNFDKVTGMGVRNGTDKDAEALFKCFRSLGFDVIVYNDCSCAKMQDLLKKAS
                     EEDHTNAACFACILLSHGEENVIYGKDGVTPIKDLTAHFRGDRCKTLLEKPKLFFIQA
                     CRGTELDDGIQADSGPINDTDANPRYKIPVEADFLFAYSTVPGYYSWRSPGRGSWFVQ
                     ALCSILEEHGKDLEIMQILTRVNDRVARHFESQSDDPHFHEKKQIPCVVSMLTKELYF
                     SQ"
BASE COUNT      768 a    515 c    616 g    702 t      1 others
ORIGIN
```

Query Match 58.2%; Score 798; DB 9; Length 2602;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1018; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
Qy    351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
Db    339 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 398

Qy    411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
Db    399 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 458

Qy    471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
Db    459 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 518

Qy    531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
Db    519 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 578

Qy    591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
Db    579 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 638

Qy    651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
Db    639 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 698

Qy    711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
Db    699 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 758

Qy    771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
Db    759 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 818

Qy    831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
Db    819 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 878

Qy    891 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 950
Db    879 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 938

Qy    951 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 1010
Db    939 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 998

Qy   1011 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 1070
Db    999 CACGGTTCCAGGCTATTACTCATGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 1058

Qy   1071 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCA 1130
Db   1059 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCA 1117

Qy   1131 GGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 1190
Db   1118 GGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 1177

Qy   1191 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAAT 1250
Db   1178 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAAT 1237

Qy   1251 AGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACAT 1310
Db   1238 AGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACAT 1297

Qy   1311 TTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTAT 1370
Db   1298 TTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTAT 1357

Qy   1371 T 1371
          |
Db   1358 T 1358
```

```
RESULT   8
HSU67206
LOCUS       HSU67206                2661 bp    mRNA    linear   PRI 03-JAN-2000
DEFINITION  Homo sapiens Lice2 alpha mRNA, complete cds.
ACCESSION   U67206
VERSION     U67206.1  GI:6649572
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2661)
  AUTHORS   Juan,T.S., McNiece,I.K., Argento,J.M., Jenkins,N.A., Gilbert,D.J.,
            Copeland,N.G. and Fletcher,F.A.
  TITLE     Identification and mapping of Casp7, a cysteine protease resembling
            CPP32 beta, interleukin-1 beta converting enzyme, and CED-3
  JOURNAL   Genomics 40 (1), 86-93 (1997)
  MEDLINE   97224489
   PUBMED   9070923
REFERENCE   2  (bases 1 to 2661)
  AUTHORS   Juan,T.S.C., McNiece,I.K., Argento,J.M., Jenkins,N.A.,
            Gilbert,D.J., Copeland,N.G. and Fletcher,F.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-AUG-1996) Developmental Hematology, Amgen, Inc., 1840
            DeHavilland Dr., Thousand Oaks, CA 91320, USA
FEATURES             Location/Qualifiers
     source          1. .2661
```

```
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             387. .1298
                     /note="like-ICE cysteine protease"
                     /codon_start=1
                     /product="Lice2 alpha"
                     /protein_id="AAF21460.1"
                     /db_xref="GI:6649573"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      740 a    569 c    642 g    709 t      1 others
ORIGIN
```

Query Match 58.2%; Score 798; DB 9; Length 2661;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1018; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
Qy     351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
Db     398 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 457

Qy     411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
Db     458 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 517

Qy     471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
Db     518 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 577

Qy     531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
Db     578 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 637

Qy     591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
Db     638 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 697

Qy     651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
Db     698 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 757

Qy     711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
Db     758 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 817

Qy     771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
Db     818 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 877

Qy     831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
Db     878 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 937

Qy     891 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 950
Db     938 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 997

Qy     951 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 1010
Db     998 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 1057

Qy    1011 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 1070
Db    1058 CACGGTTCCAGGCTATTACTCATGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 1117

Qy    1071 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCA 1130
Db    1118 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCA 1176
```

```
Qy    1131 GGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 1190
Db    1177 GGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 1236

Qy    1191 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAAT 1250
Db    1237 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAAT 1296

Qy    1251 AGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACAT 1310
Db    1297 AGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACAT 1356

Qy    1311 TTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTAT 1370
Db    1357 TTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTAT 1416

Qy    1371 T 1371
Db    1417 T 1417
```

```
RESULT   9
HSU39613
LOCUS       HSU39613                 912 bp    mRNA    linear   PRI 19-JAN-1996
DEFINITION  Human cysteine protease ICE-LAP3 mRNA, complete cds.
ACCESSION   U39613
VERSION     U39613.1  GI:1125072
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 912)
  AUTHORS   Duan,H., Chinnaiyan,A.M., Hudson,P.L., Wing,J.P., He,W.W. and
            Dixit,V.M.
  TITLE     ICE-LAP3, a novel mammalian homologue of the Caenorhabditis elegans
            cell death protein Ced-3 is activated during Fas- and tumor
            necrosis factor-induced apoptosis
  JOURNAL   J. Biol. Chem. 271 (3), 1621-1625 (1996)
  MEDLINE   96139498
REFERENCE   2  (bases 1 to 912)
  AUTHORS   Duan,H., Chinnaiyan,A.M., Hudson,P.L., Wing,J.P., He,W.W. and
            Dixit,V.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-OCT-1995) Hangjun Duan, Pathology, University of
            Michigan, 1150 W Medical Center Dr MSRBI Room 7514, Ann Arbor, MI
            48109, USA
FEATURES             Location/Qualifiers
     source          1. .912
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .912
                     /function="cysteine protease"
                     /note="member of ICE/CED-3 gene family of cysteine
                     proteases"
                     /codon_start=1
                     /product="ICE-LAP3"
                     /protein_id="AAC50346.1"
                     /db_xref="GI:1125073"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDAIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      263 a    219 c    223 g    207 t
ORIGIN
```

Query Match 52.3%; Score 717; DB 9; Length 912;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 767; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy    351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     12 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 71

Qy    411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     72 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 131

Qy    471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    132 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 191

Qy    531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    192 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 251

Qy    591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    252 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 311

Qy    651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    312 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 371

Qy    711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    372 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 431

Qy    771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    432 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 491

Qy    831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 551

Qy    891 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 950
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    552 GGCTTGCCGAGGGACCGAGCTTGATGATGCCATCCAGGCCGACTCGGGGCCCATCAATGA 611

Qy    951 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 1010
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    612 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 671

Qy   1011 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 1070
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    672 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 731

Qy   1071 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCA 1118
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    732 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCA 779
```

```
RESULT  10
HSU37449
LOCUS       HSU37449                2378 bp    mRNA    linear   PRI 13-DEC-1995
DEFINITION  Human Mch3 isoform beta (Mch3) mRNA, complete cds.
ACCESSION   U37449
VERSION     U37449.1  GI:1117848
KEYWORDS    Mch3 isoform beta.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2378)
  AUTHORS   Fernandes-Alnemri,T., Litwack,G. and Alnemri,E.S.
  TITLE     CPP32, a novel human apoptotic protein with homology to
            Caenorhabditis elegans cell death protein Ced-3 and mammalian
            interleukin-1 beta-converting enzyme
  JOURNAL   J. Biol. Chem. 269 (49), 30761-30764 (1994)
  MEDLINE   95074098
REFERENCE   2  (bases 1 to 2378)
  AUTHORS   Fernandes-Alnemri,T., Takahashi,A., Armstrong,R., Krebs,J.,
            Fritz,L., Tomaselli,K.J., Wang,L., Yu,Z., Croce,C.M., Salveson,G.,
            Earnshaw,W.C., Litwack,G. and Alnemri,E.S.
  TITLE     Mch3, a novel human apoptotic cysteine protease highly related to
            CPP32
  JOURNAL   Cancer Res. 55 (24), 6045-6052 (1995)
  MEDLINE   96105019
REFERENCE   3  (bases 1 to 2378)
  AUTHORS   Alnemri,E.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-1995) Emad S Alnemri, Pharmacology, Thomas
            Jefferson University, Jefferson Cancer Institute, 233, S. Tenth
            Street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .2378
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="T50-6A.1"
                     /cell_line="Jurkat"
                     /cell_type="T-lymphocyte"
     gene            1. .2378
                     /gene="Mch3"
     CDS             147. .908
                     /gene="Mch3"
                     /note="does not have protease activity"
                     /codon_start=1
                     /evidence=experimental
                     /product="Mch3 isoform beta"
                     /protein_id="AAC50304.1"
                     /db_xref="GI:1117849"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENMESCSVTQAGVQ
                     RRDLGRLQPPPPRLAEGPSLMMASRPTRGPSMTQMLILDTRSQWKLTSSSPIPRFQAI
                     TRGGAQEEAPGLCKPSAPSWRSTEKTWKSCRSSPG"
     polyA_site      2378
                     /gene="Mch3"
BASE COUNT      680 a    485 c    566 g    647 t
ORIGIN
```

Query Match 31.6%; Score 433; DB 9; Length 2378;
Best Local Similarity 100.0%; Pred. No. 7.9e-237;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    158 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 217

Qy    411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    218 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 277

Qy    471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    278 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 337

Qy    531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    338 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 397

Qy    591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    398 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 457

Qy    651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    458 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 517

Qy    711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    518 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 577
```

```
Qy     771 TGGAGAAGAAAAT 783
           !!!!!!!!!!!!!
Db     578 TGGAGAAGAAAAT 590


RESULT  11
AL627395
LOCUS       AL627395              155904 bp    DNA     linear   HTG 08-DEC-2001
DEFINITION  Homo sapiens chromosome 10 clone RP11-575E4, *** SEQUENCING IN
            PROGRESS ***, 3 unordered pieces.
ACCESSION   AL627395
VERSION     AL627395.11  GI:17426548
KEYWORDS    HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Lawlor,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-DEC-2001) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 8, 2001 this sequence version replaced gi:17154386.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bA575E4
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 154012 bases at least Q40
            Consensus quality: 154519 bases at least Q30
            Consensus quality: 154817 bases at least Q20
            Insert size: 155704; sum-of-contigs
            Insert size: 160371; 4.2% error; agarose-fp
            Quality coverage: 7.46x in Q20 bases; sum-of-contigs Quality
            coverage: 7.24x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1    81542: contig of 81542 bp in length
            *    81543 81642: gap of      100 bp
            *    81643    89185: contig of 7543 bp in length
            *    89186 89285: gap of      100 bp
            *    89286   155904: contig of 66619 bp in length.
FEATURES             Location/Qualifiers
     source          1. .155904
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-575E4"
                     /clone_lib="RPCI-11.2"
     misc_feature    1. .81542
                     /note="assembly_fragment:03828
                     fragment_chain:1"
     misc_feature    81643. .89185
                     /note="assembly_fragment:03912
                     fragment_chain:1"
     misc_feature    89286. .155904
                     /note="assembly_fragment:02100
                     fragment_chain:1"
BASE COUNT    44589 a  33643 c  33277 g  44195 t    200 others
ORIGIN


  Query Match             17.9%;  Score 246;  DB 2;  Length 155904;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-129;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1126 CACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTT 1185
           !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
Db  131289 CACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTT 131348

Qy    1186 CCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAG 1245
           !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
Db  131349 CCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAG 131408

Qy    1246 TCAATAGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAAT 1305
           !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
Db  131409 TCAATAGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAAT 131468

Qy    1306 CACATTTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAA 1365
           !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
Db  131469 CACATTTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAA 131528

Qy    1366 TGTATT 1371
           !!!!!!
Db  131529 TGTATT 131534


RESULT  12
AL592546/c
LOCUS       AL592546              174776 bp    DNA     linear   PRI 15-NOV-2001
DEFINITION  Human DNA sequence from clone RP11-211N11 on chromosome 10,
            complete sequence.
ACCESSION   AL592546
VERSION     AL592546.7  GI:15591647
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 174776)
  AUTHORS   Bird,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Sep 12, 2001 this sequence version replaced gi:15041956.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 10, constructed by the Sanger Centre Chromosome 10
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr10
            RP11-211N11 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
```

```
          This sequence is the entire insert of clone RP11-211N11.
FEATURES             Location/Qualifiers
     source          1. .174776
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-211N11"
                     /clone_lib="RPCI-11.1"
BASE COUNT    51535 a  35292 c  35407 g  52542 t
ORIGIN


  Query Match             17.9%;  Score 246;  DB 9;  Length 174776;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-129;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1126 CACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTT 1185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145169 CACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTT 145110

Qy    1186 CCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAG 1245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145109 CCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAG 145050

Qy    1246 TCAATAGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAAT 1305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145049 TCAATAGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAAT 144990

Qy    1306 CACATTTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAA 1365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 144989 CACATTTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAA 144930

Qy    1366 TGTATT 1371
           ||||||
Db 144929 TGTATT 144924


RESULT  13
D86353
LOCUS       D86353                   964 bp    mRNA    linear   ROD 07-FEB-1999
DEFINITION  Mouse mRNA for Mch3, complete cds.
ACCESSION   D86353
VERSION     D86353.1  GI:1945545
KEYWORDS    Mch3.
SOURCE      Mus musculus embryonal carcinoma cell cell_line:P19EC cell cDNA to
            mRNA, clone_lib:lambda cDNA library of P19ECcell.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 964)
  AUTHORS   Takashi,M., Koko,U., Takeshi,M. and Mariko,M.Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-1996) Takeshi Mukasa, National Institute of
            Neuroscience,NCNP, Division of Development and Differentiation;
            Ogawahihashi-machi 4-1-1, Kodaira, Tokyo 187, Japan
            (E-mail:momoi@ncnpj.ncnp.go.jp, Tel:81-423-42-2711,
            Fax:81-423-46-1754)
REFERENCE   2  (sites)
  AUTHORS   Mukasa,T., Urase,K., Momoi,M.Y., Kimura,I. and Momoi,T.
  TITLE     Specific expression of CPP32 in sensory neurons of mouse embryos
            and activation of CPP32 in the apoptosis induced by a withdrawal of
            NGF
  JOURNAL   Biochem. Biophys. Res. Commun. 231 (3), 770-774 (1997)
  MEDLINE   97224429
REFERENCE   3  (sites)
  AUTHORS   Mukasa,T., Khoroku,Y., Tsukahara,T., Momoi,M.Y., Kimura,I. and
            Momoi,T.
  TITLE     Wortmannin enhances CPP32-like activity during neuronal
            differentiation of P19 embryonal carcinoma cells induced by
            retinoic acid
  JOURNAL   Biochem. Biophys. Res. Commun. 232 (1), 192-197 (1997)
  MEDLINE   97236307
FEATURES             Location/Qualifiers
     source          1. .964
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /cell_line="P19EC cell"
                     /cell_type="embryonal carcinoma cell"
                     /clone_lib="lambda cDNA library of P19ECcell"
     CDS             53. .964
                     /note="similar to human Mch3alpha"
                     /codon_start=1
                     /product="Mch3"
                     /protein_id="BAA19730.1"
                     /db_xref="GI:1945546"
                     /translation="MTDDQDCAADWEKVDSSSEDGVDAKPDRSSIISSILLKKKRNAS
                     TGPRQTGRDRVPTYLYRMDFQKMGKCIIINNKNFDKATGMDVRNGTDKDAGALFKCFQ
                     NLGFEVTVHNDCSCAKMQDLLRKASEEDHSNSACFACVLLSHGEEDLIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDIDANPRNKIPVEAD
                     FLFAYSTVPGYYSWRNPGKGSWFVQALCSILNEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPRFNEKKQIPCMVSMLTKELYFSR"
BASE COUNT      246 a    260 c    263 g    195 t
ORIGIN


  Query Match             2.3%;  Score 32;  DB 10;  Length 964;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-06;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     862 TTAGAGAAACCCAAACTCTTCTTCATTCAGGC 893
           ||||||||||||||||||||||||||||||||
Db     575 TTAGAGAAACCCAAACTCTTCTTCATTCAGGC 606


RESULT  14
AF072124
LOCUS       AF072124                 1063 bp    mRNA    linear   ROD 24-MAY-2000
DEFINITION  Rattus norvegicus caspase-7 mRNA, complete cds.
ACCESSION   AF072124
VERSION     AF072124.1  GI:3243275
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1063)
  AUTHORS   Forghani,F. and Roy,S.
  TITLE     Rat caspase-7 sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1063)
  AUTHORS   Forghani,F. and Roy,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-1998) Biochemistry, Merck Frosst Centre for
            Therapeutic Research, P.O. Box 1005, Pointe-Claire-Dorval, QUE H9H
            3L1, Canada
FEATURES             Location/Qualifiers
     source          1. .1063
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /tissue_type="spleen"
     CDS             3. .914
                     /codon_start=1
                     /product="caspase-7"
                     /protein_id="AAC24011.1"
                     /db_xref="GI:3243276"
                     /translation="MTDDQDCAAELEMADSSTEDGVDAKPDRSTIISSLLWKKKKNAS
                     MCPVSTTRDRVPTYLYRMDFEKMGKCIIINNKNFDKATGMDVRNGTDKDAEALFKCFR
                     SLGFEVTVYNDCSCAKMQDLLRRASEEDHSNSACFACVLLSHGEENLIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRNPGKGSWFVQALCSILNEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPRFNEKKQIPCMVSMLTKELYFGR"
BASE COUNT      278 a    268 c    275 g    242 t
ORIGIN
```

```
  Query Match             2.3%;  Score 32;  DB 10;  Length 1063;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-06;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     628 CTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGA 659
           ||||||||||||||||||||||||||||||||
Db     291 CTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGA 322


RESULT  15
MMU67321
LOCUS       MMU67321                2006 bp    mRNA    linear   ROD 19-MAR-1997
DEFINITION  Mus musculus Lice2 cysteine protease (Lice2) mRNA, complete cds.
ACCESSION   U67321
VERSION     U67321.1  GI:1894916
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2006)
  AUTHORS   Juan,T.S., McNiece,I.K., Argento,J.M., Jenkins,N.A., Gilbert,D.J.,
            Copeland,N.G. and Fletcher,F.A.
  TITLE     Identification and mapping of Casp7, a cysteine protease resembling
            CPP32 beta, interleukin-1 beta converting enzyme, and CED-3
  JOURNAL   Genomics 40 (1), 86-93 (1997)
  MEDLINE   97224489
REFERENCE   2  (bases 1 to 2006)
  AUTHORS   Juan,T.S.C., McNiece,I.K., Argento,J.M., Jenkins,N.A.,
            Gilbert,D.J., Copeland,N.G. and Fletcher,F.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-1996) Developmental Hematology, Amgen, Inc., 1840
            DeHavilland Dr., Thousand Oaks, CA 91320, USA
FEATURES             Location/Qualifiers
     source          1. .2006
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="19"
                     /map="Between Adra2a and Aop1"
     gene            1. .2006
                     /gene="Lice2"
     CDS             474. .1496
                     /gene="Lice2"
                     /note="similar to ICE and CED-3 cysteine protease"
                     /codon_start=1
                     /product="Lice2 cysteine protease"
                     /protein_id="AAC53068.1"
                     /db_xref="GI:1894917"
                     /translation="MFTAQVFSESFTKTELLPSTLAEDGRCRGLLAAAVGTMTDDQDC
                     AAELEKVDSSSEDGVDAKPDRSSIISSILLKKKRNASAGPVRTGRDRVPTYLYRMDFQ
                     KMGKCIIINNKNFDKATGMDVRNGTDKDAGALFKCFQNLGFEVTVHNDCSCAKMQDLL
                     RKASEEDHSNSACFACVLLSHGEEDLIYGKDGVTPIKDLTAHFRGDRCKTLLEKPKLF
                     FIQACRGTELDDGIQADSGPINDIDANPRNKIPVEADFLFAYSTVPGYYSWRNPGKGS
                     WFVQALCSILNEHGKDLEIMQILTRVNDRVARHFESQSDDPRFNEKKQIPCMVSMLTK
                     ELYFSR"
BASE COUNT      543 a    490 c    462 g    511 t
ORIGIN


  Query Match             2.3%;  Score 32;  DB 10;  Length 2006;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-06;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     862 TTAGAGAAACCCAAACTCTTCTTCATTCAGGC 893
           ||||||||||||||||||||||||||||||||
Db    1107 TTAGAGAAACCCAAACTCTTCTTCATTCAGGC 1138


Search completed: August 16, 2002, 01:43:40
Job time: 6306 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 00:05:39 ; Search time 323.42 Seconds
(without alignments)
7278.120 Million cell updates/sec

Title: US-09-613-508-1
Perfect score: 1371
Sequence: 1 GCACGAGAAACTTTGCTGTG.....TAATTTCAGGAAAATGTATT 1371

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1371	100.0	1371	17	AAT15276	Human interleukin-
	2	849	61.9	2309	18	AAT66992	Apototic protease
	3	849	61.9	2309	22	AAD03914	Human caspase-7 DN
	4	833	60.8	2341	18	AAT66970	Cysteine protease
	5	768	56.0	913	20	AAX81223	Nucleotide sequenc
	6	433	31.6	2377	18	AAT66993	Cysteine protease
c	7	25	1.8	51	22	AAL26799	Human SNP oligonuc
	8	22	1.6	22	17	AAT15282	Human ICE-LAP-3 5'
c	9	22	1.6	30	18	AAT66978	Cysteine protease
c	10	22	1.6	53	17	AAT15283	Human ICE-LAP-3 3'
	11	22	1.6	72	18	AAT66973	T7-tagged CMH-1 PC
	12	21	1.5	30	18	AAT66977	Cysteine protease
	13	21	1.5	31	17	AAT15278	Human ICE-LAP-3 5'
c	14	21	1.5	31	17	AAT15279	Human ICE-LAP-3 3'
c	15	21	1.5	33	18	AAT66974	T7-tagged CMH-1 PC
c	16	20	1.5	628	22	AAS42497	Human cDNA encodin
	17	20	1.5	927	23	ABL26207	Drosophila melanog
c	18	20	1.5	1114	22	AAS26327	Human cDNA encodin
	19	20	1.5	1390	22	AAI60858	Human polynucleoti
c	20	20	1.5	1548	22	AAI59072	Human polynucleoti
c	21	20	1.5	1628	22	ABA83015	Human transcriptio
	22	20	1.5	3809	23	ABL26206	Drosophila melanog
c	23	20	1.5	4767	22	AAL02786	Human reproductive
c	24	19	1.4	19	18	AAT66994	Apoptotic protease
c	25	19	1.4	24	18	AAT66998	Apoptotic protease
c	26	19	1.4	435	19	AAV60055	Nucleic acid G75.
	27	19	1.4	494	22	AAD17530	Soybean Gly m2 pro
c	28	19	1.4	700	22	AAH92705	Human inflammatory
c	29	19	1.4	700	22	AAH92706	Human inflammatory
c	30	19	1.4	1335	21	AAC42535	Arabidopsis thalia
c	31	19	1.4	4599	16	AAT01550	Rodent (hamster/ra
c	32	19	1.4	4635	16	AAT01551	Rat sulphonylurea
c	33	19	1.4	5110	16	AAT01568	Rat sulphonylurea
	34	19	1.4	8464	21	AAZ88665	Murine NINTROX cDN
	35	19	1.4	269223	22	AAF28554	Genomic fragment #
c	36	18	1.3	18	18	AAT67000	Apoptotic protease
	37	18	1.3	24	18	AAT66999	Apoptotic protease
	38	18	1.3	232	20	AAX34087	Mycobacterium spec
	39	18	1.3	233	20	AAX34086	Mycobacterium spec
	40	18	1.3	234	20	AAX34088	Mycobacterium spec
c	41	18	1.3	273	21	AAC44656	Arabidopsis thalia
c	42	18	1.3	273	21	AAC47094	Arabidopsis thalia
	43	18	1.3	283	22	AAK76931	Human immune/haema
	44	18	1.3	290	22	AAK60102	Human immune/haema
	45	18	1.3	291	21	AAZ90476	Cancer specific ge

ALIGNMENTS

RESULT 1
AAT15276
ID AAT15276 standard; cDNA; 1371 BP.
XX
AC AAT15276;
XX
DT 27-OCT-1996 (first entry)
XX
DE Human interleukin-I-converting enzyme-like apoptosis protease-3 cDNA.
XX
KW ICE-LAP-3; interleukin-I-converting enzyme-like apoptosis protease;
KW enzyme; Alzheimer's disease; Parkinson's disease; septic shock;
KW head injury; rheumatoid arthritis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 115..1137
FT /*tag= a
XX
PN WO9613603-A1.
XX
PD 09-MAY-1996.
XX
PF 06-JUN-1995; 95WO-US07235.
XX
PR 01-NOV-1994; 94US-0334251.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Hastings GA, He W, Hudson PL, Rosen CA;
XX

```
DR   WPI; 1996-239509/24.
DR   P-PSDB; AAR95830.
XX
PT   Human ICE-LAP-3 and -4 DNA and protein  - useful in the diagnosis
PT   and treatment of Alzheimer's disease, Parkinson's disease,
PT   rheumatoid arthritis, septic shock and head injury
XX
PS   Claim 1; Page 91-92; 67pp; English.
XX
CC   This sequence was derived from a cDNA library from human endometrial
CC   tumuor and may be used for the recombinant production of the ICE-LAP-
CC   3 protein e.g. in COS cells.  The encoded protein may be used
CC   therapeutically, e.g. as an antitumor or antiviral agent and to
CC   control embryonic development and tissue homeostasis.  The protein
CC   can also be used to treat immunosuppression disorders, such as AIDS,
CC   by targeting virus infected cells for cell death.  The DNA may find
CC   use in gene therapy applications.
XX
SQ   Sequence 1371 BP; 355 A; 357 C; 354 G; 305 T; 0 other;
```

```
  Query Match             100.0%;  Score 1371;  DB 17;  Length 1371;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1371;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCACGAGAAACTTTGCTGTGCGCGTTCTCCCGCGCGCGGGCTCAACTTTGTAGAGCGAGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 gcacgagaaactttgctgtgcgcgttctcccgcgcgcgggctcaactttgtagagcgagg 60

Qy      61 GGCCAACTTGGCAGAGCGCGCGGCCAGCTTTGCAGAGAGCGCCCTCCAGGGACTATGCGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ggccaacttggcagagcgcgcggccagctttgcagagagcgccctccagggactatgcgt 120

Qy     121 GCGGGGACACGGGTCGCTTTGGGCTCTTCCACCCCTGCGGAGCGCACTACCCCGAGCCAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gcggggacacgggtcgctttgggctcttccaccccgcggagcgcactaccccgagccag 180

Qy     181 GGGCGGTGCAAGCCCCGCCCGGCCCTACCCAGGGCGGCTCCTCCCTCCGCAGCGCCGAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 gggcggtgcaagccccgcccggccctacccagggcggctcctccctccgcagcgccgaga 240

Qy     241 CTTTTAGTTTCGCTTTCGCTAAAGGGGCCCCAGACCCTTGCTGCGGAGCGACGGAGAGAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 cttttagtttcgctttcgctaaaggggccccagacccttgctgcggagcgacggagagag 300

Qy     301 ACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGGGCTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 actgtgccagtcccagccgccctaccgccgtgggaacgatggcagatgattcagggctgt 360

Qy     361 ATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 attgaagagcaggggttgaggattcagcaaatgaagattcagtggatgctaagccagac 420

Qy     421 CGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 cggtcctcgtttgtaccgtccctcttcagtaagaagaagaaaaatgtcaccatgcgatcc 480

Qy     481 ATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 atcaagaccacccgggaccgagtgcctacatatcagtacaacatgaattttgaaaagctg 540

Qy     541 GGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 ggcaaatgcatcataataaacaacaagaactttgataaagtgacaggtatgggcgttcga 600

Qy     601 AACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 aacggaacagacaaagatgccgaggcgctcttcaagtgcttccgaagcctgggttttgac 660

Qy     661 GTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 gtgattgtctataatgactgctcttgtgccaagatgcaagatctgcttaaaaaagcttct 720

Qy     721 GAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 gaagaggaccatacaaatgccgcctgcttcgcctgcatcctcttaagccatggagaagaa 780

Qy     781 AATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 aatgtaatttatgggaaagatggtgtcacaccaataaaggatttgacagcccactttagg 840

Qy     841 GGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 ggggatagatgcaaaaccctttagagaaacccaaactcttcttcattcaggcttgccga 900

Qy     901 GGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 gggaccgagcttgatgatggcatccaggccgactcggggcccatcaatgacacagatgct 960

Qy     961 AATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 aatcctcgatacaagatcccagtggaagctgacttcctcttcgcctattccacggttcca 1020

Qy    1021 GGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 ggctattactcgtggaggagcccaggaagaggctcctggtttgtgcaagccctctgctcc 1080

Qy    1081 ATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 atcctggaggagcacggaaaagacctggaaatcatgcaaatcctccaccagggtgaatga 1140

Qy    1141 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 cagagttgccaggcactttgagtctcagtctgatgacccacacttccatgagaagaagca 1200

Qy    1201 GATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 gatccctgtgtggtctccatgctcaccaaggaactctacttcagtcaatagccatatca 1260

Qy    1261 GGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 ggggtacattctagctgagaagcaatgggtcactcattaatgaatcacatttttttatgc 1320

Qy    1321 TCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 tcttgaaatattcagaaattctccaggattttaatttcaggaaaatgtatt 1371
```

```
RESULT   2
AAT66992
ID   AAT66992 standard; cDNA; 2309 BP.
XX
AC   AAT66992;
XX
DT   05-AUG-1997  (first entry)
XX
DE   Apototic protease Mch3-alpha cDNA.
XX
KW   Mch3-alpha; cysteine protease; apoptosis; AIDS; ischaemia;
KW   neurodegenerative disease; gene therapy; diagnosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             44..955
FT                   /*tag= a
FT   intron`         595..596
FT                   /*tag= b
FT                   /note= "site of intron in genomic DNA"
XX
PN   WO9718313-A1.
XX
PD   22-MAY-1997.
XX
PF   12-NOV-1996;   96WO-US18118.
XX
```

```
PR   13-NOV-1995;   95US-0556627.
XX
PA   (IDUN-) IDUN PHARM INC.
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES,  Armstrong R,  Fernandes-Alnemri T,  Litwack G;
PI   Tomaselli K;
XX
DR   WPI; 1997-289289/26.
DR   P-PSDB; AAW15262.
XX
PT   New gene encoding Mch3, a cysteine protease that regulates apoptosis
PT   - for treating human diseases associated with apoptosis, and
PT   screening for antagonists and agonists of Mch3
XX
PS   Claim 2; Fig 1; 52pp; English.
XX
CC   A cDNA clone (AAT66992) codes for Mch3-alpha (AAW15262), an apoptotic
CC   protein and novel member of the interleukin-1-beta converting
CC   enzyme (ICE) family of cysteine proteases.  It was obtd. by
CC   screening a human database of expressed sequence tags for sequences
CC   showing homology to ICE family members, and use of a partial cDNA
CC   clone with high homology to CPP32 and Ced-3 to screen a Jurkat cDNA
CC   library.  A second cDNA clone, Mch3-beta (AAT66993), contains a
CC   deletion and insertion corresponding to nucleotides 488-592 of the
CC   Mch3-alpha sequence.  Mch3 nucleic acids can be used to produce
CC   recombinant Mch3 polypeptides useful for treating diseases
CC   associated with cell death, such as AIDS, ischaemic injury,
CC   neurodegenerative diseases, etc.  They can also be used in gene
CC   therapy to regulate apoptosis.
XX
SQ   Sequence 2309 BP; 680 A; 453 C; 531 G; 645 T; 0 other;
```

```
  Query Match             61.9%;  Score 849;  DB 18;  Length 2309;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1019;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy     351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55 tcagggctgtattgaagagcaggggttgaggattcagcaaatgaagattcagtggatgc 114

Qy     411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115 taagccagaccggtcctcgtttgtaccgtccctcttcagtaagaagaagaaaaatgtcac 174

Qy     471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     175 catgcgatccatcaagaccacccgggaccgagtgcctacatatcagtacaacatgaattt 234

Qy     531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     235 tgaaaagctgggcaaatgcatcataataaacaacaagaactttgataaagtgacaggtat 294

Qy     591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     295 gggcgttcgaaacggaacagacaaagatgccgaggcgctcttcaagtgcttccgaagcct 354

Qy     651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     355 gggttttgacgtgattgtctataatgactgctcttgtgccaagatgcaagatctgcttaa 414

Qy     711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     415 aaaagcttctgaagaggaccatacaaatgccgcctgcttcgcctgcatcctcttaagcca 474

Qy     771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     475 tggagaagaaaatgtaatttatgggaaagatggtgtcacaccaataaaggatttgacagc 534

Qy     831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     535 ccactttaggggggatagatgcaaaacccttttagagaaacccaaactcttcttcattca 594

Qy     891 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 950
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     595 ggcttgccgagggaccgagcttgatgatggcatccaggccgactcggggcccatcaatga 654

Qy     951 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 1010
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     655 cacagatgctaatcctcgatacaagatcccagtggaagctgacttcctcttcgcctattc 714

Qy    1011 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 1070
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     715 cacggttccaggctattactcgtggaggagcccaggaagaggctcctggtttgtgcaagc 774

Qy    1071 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCA 1130
           |||||||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db     775 cctctgctccatcctggaggagcacggaaaagacctggaaatcatgcagatcct-cacca 833

Qy    1131 GGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 1190
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     834 gggtgaatgacagagttgccaggcactttgagtctcagtctgatgacccacacttccatg 893

Qy    1191 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAAT 1250
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     894 agaagaagcagatcccctgtgtggtctccatgctcaccaaggaactctacttcagtcaat 953

Qy    1251 AGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACAT 1310
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     954 agccatatcaggggtacattctagctgagaagcaatgggtcactcattaatgaatcacat 1013

Qy    1311 TTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTAT 1370
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1014 ttttttatgctcttgaaatattcagaaattctccaggattttaatttcaggaaaatgtat 1073

Qy    1371 T 1371
           |
Db    1074 t 1074
```

```
RESULT   3
AAD03914
ID   AAD03914 standard; DNA; 2309 BP.
XX
AC   AAD03914;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Human caspase-7 DNA.
XX
KW   Human; caspase-7; interleukin-1 converting enzyme; ICE-LAP3; CHM-1; Mch3;
KW   cysteine protease; apoptosis; caspase expression cassette; metastasis;
KW   tumour; cathepsin B; urokinase; proliferation; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             44..955
FT                   /*tag= a
FT                   /product= "Human caspase-7"
XX
PN   WO200129232-A2.
XX
PD   26-APR-2001.
XX
PF   19-OCT-2000; 2000WO-US28941.
XX
PR   20-OCT-1999;   99US-0160559.
PR   14-AUG-2000; 2000US-0225564.
XX
PA   (SCIO-) SCIOS INC.
XX
PI   Cordell B,  Li Y;
XX
DR   WPI; 2001-290920/30.
```

```
DR   P-PSDB; AAE00604.
XX
PT   Novel fusion polypeptide comprising first and second caspase subunit
PT   separated by cleavage site not associated in nature with caspase
PT   subunit, useful for cloning gene encoding enzymes involved in
PT   proteolytic cleavage  -
XX
PS   Disclosure; Fig 13; 116pp; English.
XX
CC   The present sequence is a DNA encoding human Caspase-7 also known as
CC   interleukin-1 converting enzyme (ICE) LAP3, CHM-1 and Mhc3. Caspases are
CC   a family of cysteine proteases, that participate in the initiation and
CC   execution of apoptosis. Caspases exist as pro-enzymes, activated by
CC   cleavage into a large and small subunit, occurring after specific
CC   aspartic acid residues within the pro-enzyme sequence.
CC   The present invention relates to a method for functional cloning of genes
CC   encoding proteins or enzymes involved in proteolytic cleavage. The
CC   invention is based on the use of caspase expression cassettes comprising
CC   the coding sequence of a proteolytic cleavage site flanked by sequences
CC   encoding two caspase subunits. A fusion polypeptide comprising a first
CC   and a second caspase subunit, separated by a cleavage site not associated
CC   in nature, is useful for cloning gene encoding enzymes involved in
CC   proteolytic cleavage. An expression cassette containing fusion
CC   polypeptide is used to identify a mutant cell line deficient in an
CC   enzyme of interest and is also useful for diagnosis and suppression of
CC   proliferation or metastases of a tumour cell characterised by
CC   overexpression of a polypeptide (e.g. Cathepsin B or urokinase,
CC   selectively expressed in the tumour cells). DNA encoding fusion
CC   polypeptide is used in gene therapy.
XX
SQ   Sequence 2309 BP; 680 A; 453 C; 531 G; 645 T; 0 other;

  Query Match             61.9%;  Score 849;  DB 22;  Length 2309;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1019;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy     351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55 tcagggctgtattgaagagcaggggttgaggattcagcaaatgaagattcagtggatgc 114

Qy     411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115 taagccagaccggtcctcgtttgtaccgtccctcttcagtaagaagaagaaaaatgtcac 174

Qy     471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     175 catgcgatccatcaagaccacccgggaccgagtgcctacatatcagtacaacatgaatttt 234

Qy     531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     235 tgaaaagctgggcaaatgcatcataataaacaacaagaactttgataaagtgacaggtat 294

Qy     591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     295 gggcgttcgaaacggaacagacaaagatgccgaggcgctcttcaagtgcttccgaagcct 354

Qy     651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     355 gggttttgacgtgattgtctataatgactgctcttgtgccaagatgcaagatctgcttaa 414

Qy     711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     415 aaaagcttctgaagaggaccatacaaatgccgcctgcttcgcctgcatcctcttaagcca 474

Qy     771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     475 tggagaagaaaatgtaatttatgggaaagatggtgtcacaccaataaaggatttgacagc 534

Qy     831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     535 ccactttaggggggatagatgcaaaaccctttttagagaaacccaaactcttcttcattca 594

Qy     891 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 950
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     595 ggcttgccgagggaccgagcttgatgatggcatccaggccgactcggggcccatcaatga 654

Qy     951 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 1010
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     655 cacagatgctaatcctcgatacaagatcccagtggaagctgacttcctcttcgcctattc 714

Qy    1011 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 1070
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     715 cacggttccaggctattactcgtggaggagcccaggaagaggctcctggtttgtgcaagc 774

Qy    1071 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCA 1130
           ||||||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db     775 cctctgctccatcctggaggagcacggaaaagacctggaaatcatgcagatcct-cacca 833

Qy    1131 GGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 1190
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     834 gggtgaatgacagagttgccaggcactttgagtctcagtctgatgacccacacttccatg 893

Qy    1191 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAAT 1250
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     894 agaagaagcagatcccctgtgtggtctccatgctcaccaaggaactctacttcagtcaat 953

Qy    1251 AGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACAT 1310
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     954 agccatatcaggggtacattctagctgagaagcaatgggtcactcattaatgaatcacat 1013

Qy    1311 TTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTAT 1370
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1014 ttttttatgctcttgaaatattcagaaattctccaggattttaatttcaggaaaatgtat 1073

Qy    1371 T 1371
           |
Db    1074 t 1074


RESULT   4
AAT66970
ID   AAT66970 standard; cDNA; 2341 BP.
XX
AC   AAT66970;
XX
DT   21-JUL-1997  (first entry)
XX
DE   Cysteine protease CMH-1 cDNA.
XX
KW   Cysteine protease; CMH-1; CPP32/Mch2 homologue-1; apoptosis;
KW   programmed cell death; cancer; neurodegenerative disease;
KW   autoimmune disease; gene therapy; diagnosis; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             54..965
FT                   /*tag= a
FT                   /product= CMH-1 proenzyme
FT                   /note= "bases 123-962 and 138-962 encode active
FT                           active CMH-1 (Claim 2)"
XX
PN   WO9716552-A1.
XX
PD   09-MAY-1997.
XX
PF   01-NOV-1996;   96WO-US17431.
XX
PR   16-NOV-1995;   95US-0558733.
PR   03-NOV-1995;   95US-0007211.
PR   06-NOV-1995;   95US-0007251.
XX
PA   (VERT-) VERTEX PHARM INC.
XX
PI   Lippke JA,  Su M;
```

```
XX
DR   WPI; 1997-272121/24.
DR   P-PSDB; AAW15247.
XX
PT   DNA encoding active, activatable or inactive cysteine protease CMH-1
PT   - useful in gene therapy for promoting and inhibiting apoptosis and
PT   for diagnosing cells with potential for apoptosis
XX
PS   Claim 1; Page 30-32; 48pp; English.
XX
CC   A cDNA clone (T66970) includes to coding sequence for the proenzyme
CC   (W15247) of CMH-1, a human cysteine protease involved in apoptosis.
CC   To obtain the clone, the GenBank Expressed Sequence Tag (EST)
CC   database was screened for clones encoding polypeptides similar to
CC   interleukin-1b converting enzyme.  Partial cDNA clone T50828 was
CC   identified.  Partial cDNA clone 72778 corresponding to this EST was
CC   then obtd.  A spleen cDNA pool was subjected to PCR using primers
CC   (T66971-92) based on a 5' adaptor and EST T50828.  A PCR fragment
CC   contg. the 5' coding region was ligated to the clone 72778 3' region
CC   to obtain the full-length CHM-1 sequence.  CHM-1 (i.e. CPP32/Mch2
CC   homologue-1) DNA can be used in gene therapy to promote apoptosis,
CC   esp. in cancer cells, and to modulate patient cholesterol levels,
CC   as well as to diagnose cells with potential for apoptosis.
CC   Antisense sequences or sequences encoding inactive CHM-1 mutants
CC   can be used to inhibit apoptosis, e.g. to prevent neurological
CC   disease.
XX
SQ   Sequence 2341 BP; 702 A; 455 C; 536 G; 648 T; 0 other;

  Query Match             60.8%;  Score 833;  DB 18;  Length 2341;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1003;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy     367 GAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCC 426
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      81 gagcaggggttgaggattcagcaaatgaagattcagtggatgctaagccagaccggtcc 140

Qy     427 TCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAG 486
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     141 tcgtttgtaccgtccctcttcagtaagaagaagaaaaatgtcaccatgcgatccatcaag 200

Qy     487 ACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAA 546
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     201 accacccgggaccgagtgcctacatatcagtacaacatgaattttgaaaagctgggcaaa 260

Qy     547 TGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGA 606
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     261 tgcatcataataaacaacaagaactttgataaagtgacaggtatgggcgttcgaaacgga 320

Qy     607 ACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATT 666
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     321 acagacaaagatgccgaggcgctcttcaagtgcttccgaagcctgggttttgacgtgatt 380

Qy     667 GTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAG 726
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     381 gtctataatgactgctcttgtgccaagatgcaagatctgcttaaaaaagcttctgaagag 440

Qy     727 GACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTA 786
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     441 gaccatacaaatgccgcctgcttcgcctgcatcctcttaagccatggagaagaaaatgta 500

Qy     787 ATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGAT 846
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     501 atttatgggaaagatggtgtcacaccaataaaggatttgacagcccactttaggggggat 560

Qy     847 AGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACC 906
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     561 agatgcaaaaccctttagagaaacccaaactcttcttcattcaggcttgccgagggacc 620

Qy     907 GAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCT 966
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     621 gagcttgatgatggcatccaggccgactcggggcccatcaatgacacagatgctaatcct 680

Qy     967 CGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTAT 1026
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     681 cgatacaagatcccagtggaagctgacttcctcttcgcctattccacggttccaggctat 740

Qy    1027 TACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTG 1086
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     741 tactcgtggaggagcccaggaagaggctcctggtttgtgcaagccctctgctccatcctg 800

Qy    1087 GAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGACAGAGT 1146
           |||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||||
Db     801 gaggagcacggaaaagacctggaaatcatgcagatcct-caccagggtgaatgacagagt 859

Qy    1147 TGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGATCCC 1206
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     860 tgccaggcactttgagtctcagtctgatgacccacacttccatgagaagaagcagatccc 919

Qy    1207 CTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCAGGGGTA 1266
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     920 ctgtgtggtctccatgctcaccaaggaactctacttcagtcaatagccatatcaggggta 979

Qy    1267 CATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGCTCTTGA 1326
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     980 cattctagctgagaagcaatgggtcactcattaatgaatcacatttttttatgctcttga 1039

Qy    1327 AATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
           |||||||||||||||||||||||||||||||||||||||||||||
Db    1040 aatattcagaaattctccaggattttaatttcaggaaaatgtatt 1084


RESULT   5
AAX81223
ID   AAX81223 standard; DNA; 913 BP.
XX
AC   AAX81223;
XX
DT   10-SEP-1999  (first entry)
XX
DE   Nucleotide sequence of caspase-7 (Mch3).
XX
KW   Rev-caspase; cysteine protease; zymogen; caspase; cancer; gene therapy;
KW   autoimmune disease; caspase-mediated apoptosis; neurodegenerative;
KW   tumour cell; myocardial infarction; human; ds.
XX
OS   Homo sapiens.
XX
PN   WO9935277-A2.
XX
PD   15-JUL-1999.
XX
PF   11-JAN-1999;   99WO-US00632.
XX
PR   09-JAN-1998;   98US-0070987.
XX
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES;
XX
DR   WPI; 1999-419353/35.
DR   P-PSDB; AAY21721.
XX
PT   New isolated nucleic acid molecule encoding a rev-caspase  - used
PT   for screening and identifying inhibitors or enhancers for treating
PT   cancer or autoimmune disease
XX
PS   Disclosure; Fig 17A-B; 74pp; English.
XX
CC   The invention relates to nucleic acid molecules encoding rev-caspases.
CC   Rev-caspases are cysteine proteases that specifically cleave proteins
CC   after Asp residues and is expressed as a zymogen, in which the small
CC   subunit is N-terminal to a large subunit. A gene delivery vehicle
CC   comprising a rev-caspase coding sequence is useful for the treatment of
CC   cancer, where the gene delivery vehicle is internalised by tumour cells.
```

```
CC   The gene delivery vehicle can also be used to treat autoimmune diseases.
CC   Cells transfected with a rev-caspase expressing vector can be used in
CC   identification of inhibitors or enhancers of caspase-mediated apoptosis.
CC   In vitro translated rev-caspase can be used to identify an inhibitor or
CC   enhancer of caspase processing activity. Caspase inhibitors are useful
CC   for treating neurodegenerative diseases as well as for inhibiting
CC   apoptosis in the heart following myocardial infarction. Sequences
CC   AAX81217 -AAX81226 represent human caspase genes encoding caspase 1-10
CC   gene products (AAY21715-Y21724).
XX
SQ   Sequence 913 BP; 263 A; 218 C; 224 G; 207 T; 1 other;
```

```
  Query Match             56.0%;  Score 768;  DB 20;  Length 913;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  768;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12 tcagggctgtattgaagagcaggggttgaggattcagcaaatgaagattcagtggatgc 71

Qy     411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72 taagccagaccggtcctcgtttgtaccgtccctcttcagtaagaagaagaaaaatgtcac 131

Qy     471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     132 catgcgatccatcaagaccacccgggaccgagtgcctacatatcagtacaacatgaattt 191

Qy     531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     192 tgaaaagctgggcaaatgcatcataataaacaacaagaactttgataaagtgacaggtat 251

Qy     591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     252 gggcgttcgaaacggaacagacaaagatgccgaggcgctcttcaagtgcttccgaagcct 311

Qy     651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     312 gggttttgacgtgattgtctataatgactgctcttgtgccaagatgcaagatctgcttaa 371

Qy     711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     372 aaaagcttctgaagaggaccatacaaatgccgcctgcttcgcctgcatcctcttaagcca 431

Qy     771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     432 tggagaagaaaatgtaatttatgggaaagatggtgtcacaccaataaaggatttgacagc 491

Qy     831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     492 ccactttaggggggatagatgcaaaaccctttagagaaacccaaactcttcttcattca 551

Qy     891 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 950
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     552 ggcttgccgagggaccgagcttgatgatggcatccaggccgactcggggcccatcaatga 611

Qy     951 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 1010
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     612 cacagatgctaatcctcgatacaagatcccagtggaagctgacttcctcttcgcctattc 671

Qy    1011 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 1070
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     672 cacggttccaggctattactcgtggaggagcccaggaagaggctcctggtttgtgcaagc 731

Qy    1071 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCA 1118
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db     732 cctctgctccatcctggaggagcacggaaaagacctggaaatcatgca 779
```

```
RESULT   6
AAT66993
ID   AAT66993 standard; cDNA; 2377 BP.
XX
AC   AAT66993;
XX
DT   05-AUG-1997  (first entry)
XX
DE   Cysteine protease Mch3-beta cDNA.
XX
KW   Mch3-beta; cysteine protease; apoptosis; AIDS; ischaemia;
KW   neurodegenerative disease; gene therapy; diagnosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             146..839
FT                   /*tag= a
XX
PN   WO9718313-A1.
XX
PD   22-MAY-1997.
XX
PF   12-NOV-1996;   96WO-US18118.
XX
PR   13-NOV-1995;   95US-0556627.
XX
PA   (IDUN-) IDUN PHARM INC.
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES,  Armstrong R,  Fernandes-Alnemri T,  Litwack G;
PI   Tomaselli K;
XX
DR   WPI; 1997-289289/26.
DR   P-PSDB; AAW15263.
XX
PT   New gene encoding Mch3, a cysteine protease that regulates apoptosis
PT   - for treating human diseases associated with apoptosis, and
PT   screening for antagonists and agonists of Mch3
XX
PS   Claim 2; Fig 1; 52pp; English.
XX
CC   A cDNA clone (AAT66993) codes for Mch3-beta (AAW15263), a novel member
CC   of the interleukin-1-beta converting enzyme (ICE) family of
CC   cysteine proteases.  It was obtd. by screening a human database of
CC   expressed sequence tags for sequences showing homology to ICE
CC   family members, and use of a partial cDNA clone with high homology
CC   to CPP32 and Ced-3 to screen a Jurkat cDNA library.  Mch3-beta cDNA
CC   contains a deletion and insertion corresponding to nucleotides
CC   488-592 of Mch3-alpha cDNA (AAT66992) and also has a longer 3'
CC   nontranslated sequence.  Mch3-beta lacks the active site of Mch3-
CC   alpha.  Mch3 nucleic acids can be used to produce recombinant Mch3
CC   polypeptides useful for treating diseases associated with cell
CC   death, such as AIDS, ischaemic injury, neurodegenerative diseases,
CC   etc.  They can also be used in gene therapy to regulate apoptosis.
XX
SQ   Sequence 2377 BP; 680 A; 485 C; 565 G; 647 T; 0 other;
```

```
  Query Match             31.6%;  Score 433;  DB 18;  Length 2377;
  Best Local Similarity   100.0%;  Pred. No. 7.9e-198;
  Matches  433;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     157 tcagggctgtattgaagagcaggggttgaggattcagcaaatgaagattcagtggatgc 216

Qy     411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     217 taagccagaccggtcctcgtttgtaccgtccctcttcagtaagaagaagaaaaatgtcac 276

Qy     471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     277 catgcgatccatcaagaccacccgggaccgagtgcctacatatcagtacaacatgaattt 336

Qy     531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    337 tgaaaagctgggcaaatgcatcataataaacaacaagaactttgataaagtgacaggtat 396

Qy    591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    397 gggcgttcgaaacggaacagacaaagatgccgaggcgctcttcaagtgcttccgaagcct 456

Qy    651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    457 gggttttgacgtgattgtctataatgactgctcttgtgccaagatgcaagatctgcttaa 516

Qy    711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    517 aaaagcttctgaagaggaccatacaaatgccgcctgcttcgcctgcatcctcttaagcca 576

Qy    771 TGGAGAAGAAAAT 783
          |||||||||||||
Db    577 tggagaagaaaat 589
```

RESULT 7
AAL26799/c
ID AAL26799 standard; DNA; 51 BP.
XX
AC AAL26799;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #7.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 1402; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 51 BP; 11 A; 14 C; 14 G; 12 T; 0 other;

```
  Query Match             1.8%;  Score 25;  DB 22;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 0.072;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1007 ATTCCACGGTTCCAGGCTATTACTC 1031
          |||||||||||||||||||||||||
Db     51 ATTCCACGGTTCCAGGCTATTACTC 27
```

RESULT 8
AAT15282
ID AAT15282 standard; cDNA; 22 BP.
XX
AC AAT15282;
XX
DT 28-OCT-1996 (first entry)
XX
DE Human ICE-LAP-3 5' oligonucleotide DNA primer.
XX
KW ICE-LAP-3; interleukin-I-converting enzyme-like apoptosis protease;
KW enzyme; Alzheimer's disease; Parkinson's disease; septic shock;
KW head injury; rheumatoid arthritis; PCR; polymerae chain reaction;
KW DNA primer; plasmid ICE-LAP-3HA; ss.
XX
OS Synthetic.
XX
PN WO9613603-A1.
XX
PD 09-MAY-1996.
XX
PF 06-JUN-1995; 95WO-US07235.
XX
PR 01-NOV-1994; 94US-0334251.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Hastings GA, He W, Hudson PL, Rosen CA;
XX
DR WPI; 1996-239509/24.
XX
PT Human ICE-LAP-3 and -4 DNA and protein - useful in the diagnosis
PT and treatment of Alzheimer's disease, Parkinson's disease,
PT rheumatoid arthritis, septic shock and head injury
XX
PS Disclosure; Page 34; 67pp; English.
XX
CC This 5' oligonucleotide DNA primer was used in a strategy for the
CC construction of plasmid ICE-LAP-3HA, and was used specifically for the
CC construction of the ICE-LAP-3 (ATCC 75875) DNA sequence (in
CC conjunction with a 3' DNA primer (AAT15283)). This 5' DNA primer
CC contains the ICE-LAP-3 translational initiation site ATG followed
CC by 5 nucleotides of ICE-LAP-3 coding sequence starting from the
CC initiation codon.
XX
SQ Sequence 22 BP; 4 A; 5 C; 10 G; 3 T; 0 other;

```
  Query Match             1.6%;  Score 22;  DB 17;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    111 GACTATGCGTGCGGGGACACGG 132
          ||||||||||||||||||||||
Db      1 gactatgcgtgcggggacacgg 22
```

```
RESULT   9
AAT66978/c
ID   AAT66978 standard; DNA; 30 BP.
XX
AC   AAT66978;
XX
DT   21-JUL-1997  (first entry)
XX
DE   Cysteine protease CMH-1 PCR primer.
XX
KW   Cysteine protease; CMH-1; CPP32/Mch2 homologue-1; apoptosis;
KW   programmed cell death; cancer; neurodegenerative disease;
KW   autoimmune disease; gene therapy; diagnosis; primer; PCR;
KW   polymerase chain reaction; ss.
XX
OS   Synthetic.
XX
PN   WO9716552-A1.
XX
PD   09-MAY-1997.
XX
PF   01-NOV-1996;   96WO-US17431.
XX
PR   16-NOV-1995;   95US-0558733.
PR   03-NOV-1995;   95US-0007211.
PR   06-NOV-1995;   95US-0007251.
XX
PA   (VERT-) VERTEX PHARM INC.
XX
PI   Lippke JA,  Su M;
XX
DR   WPI; 1997-272121/24.
XX
PT   DNA encoding active, activatable or inactive cysteine protease CMH-1
PT   - useful in gene therapy for promoting and inhibiting apoptosis and
PT   for diagnosing cells with potential for apoptosis
XX
PS   Example 3; Page 25; 48pp; English.
XX
CC   PCR primers (AAT66977-78) were designed to introduce in-frame XhoI
CC   sites to the 5' and 3' ends of human cysteine protease cmh-1
CC   cDNA (see also AAT66970).  The PCR product XhoI fragment was ligated
CC   into pEt-15b inducible E. coli expression vector.  The resulting
CC   plasmid encoded a polypeptide comprising a 23-residue N-terminal
CC   peptide (AAW15248) contg. a hexahistidine tag and thrombin cleavage
CC   site, fused to Ala24 of CHM-1.  Hexahistidine-tagged CMH-1 (Ala24-
CC   Gln303) was produced in E. coli BM21 (DE3) transformants.
XX
SQ   Sequence 30 BP; 7 A; 5 C; 10 G; 8 T; 0 other;
```

```
  Query Match             1.6%;  Score 22;  DB 18;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1232 GAACTCTACTTCAGTCAATAGC 1253
           ||||||||||||||||||||||
Db      30 GAACTCTACTTCAGTCAATAGC 9
```

```
RESULT  10
AAT15283/c
ID   AAT15283 standard; cDNA; 53 BP.
XX
AC   AAT15283;
XX
DT   28-OCT-1996  (first entry)
XX
DE   Human ICE-LAP-3 3' oligonucleotide DNA primer.
XX
KW   ICE-LAP-3; interleukin-I-converting enzyme-like apoptosis protease;
KW   enzyme; Alzheimer's disease; Parkinson's disease; septic shock;
KW   head injury; rheumatoid arthritis; PCR; polymerae chain reaction;
KW   DNA primer; plasmid ICE-LAP-3HA; ss.
XX
OS   Synthetic.
XX
PN   WO9613603-A1.
XX
PD   09-MAY-1996.
XX
PF   06-JUN-1995;   95WO-US07235.
XX
PR   01-NOV-1994;   94US-0334251.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Hastings GA,  He W,  Hudson PL,  Rosen CA;
XX
DR   WPI; 1996-239509/24.
XX
PT   Human ICE-LAP-3 and -4 DNA and protein  - useful in the diagnosis
PT   and treatment of Alzheimer's disease, Parkinson's disease,
PT   rheumatoid arthritis, septic shock and head injury
XX
PS   Disclosure; Page 34; 67pp; English.
XX
CC   This 3' oligonucleotide DNA primer was used in a strategy for the
CC   construction of plasmid ICE-LAP-3HA, and was used specifically for the
CC   construction of the ICE-LAP-3 (ATCC 75875) DNA sequence (in
CC   conjunction with a 5' DNA primer (AAT15282)).  This 3' DNA primer
CC   contains the ICE-LAP-3 translation stop codon, HA tag and the last
CC   21 nucleotides of the ICE-LAP-3 coding sequence (not including the
CC   stop codon.
XX
SQ   Sequence 53 BP; 11 A; 9 C; 18 G; 15 T; 0 other;
```

```
  Query Match             1.6%;  Score 22;  DB 17;  Length 53;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1117 CAAATCCTCCACCAGGGTGAAT 1138
           ||||||||||||||||||||||
Db      53 CAAATCCTCCACCAGGGTGAAT 32
```

```
RESULT  11
AAT66973
ID   AAT66973 standard; DNA; 72 BP.
XX
AC   AAT66973;
XX
DT   21-JUL-1997  (first entry)
XX
DE   T7-tagged CMH-1 PCR primer.
XX
KW   Cysteine protease; CMH-1; CPP32/Mch2 homologue-1; apoptosis;
KW   programmed cell death; cancer; neurodegenerative disease;
KW   autoimmune disease; gene therapy; diagnosis; primer; PCR;
KW   polymerase chain reaction; ss.
XX
OS   Synthetic.
XX
PN   WO9716552-A1.
XX
PD   09-MAY-1997.
XX
PF   01-NOV-1996;   96WO-US17431.
XX
PR   16-NOV-1995;   95US-0558733.
PR   03-NOV-1995;   95US-0007211.
PR   06-NOV-1995;   95US-0007251.
XX
```

PA (VERT-) VERTEX PHARM INC.
XX
PI Lippke JA, Su M;
XX
DR WPI; 1997-272121/24.
XX
PT DNA encoding active, activatable or inactive cysteine protease CMH-1
PT - useful in gene therapy for promoting and inhibiting apoptosis and
PT for diagnosing cells with potential for apoptosis
XX
PS Example 2; Page 23; 48pp; English.
XX
CC PCR primers (AAT66973-74) were used to produce a cDNA fragment
CC encoding an N-terminally T7-tagged cysteine protein CMH-1 lacking
CC the first 23 amino acid residues of the proenzyme (see also
CC AAW15247). Amplified DNA was subcloned into an SV40-based expression
CC vector under control of the SR-alpha promoter. COS cells
CC transiently transfected with the vector resulted in high
CC expression of a truncated CMH-1 protein. Over-expression of
CC the protein resulted in cell apoptosis.
XX
SQ Sequence 72 BP; 16 A; 19 C; 23 G; 14 T; 0 other;

Query Match 1.6%; Score 22; DB 18; Length 72;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 408 TGCTAAGCCAGACCGGTCCTCG 429
||||||||||||||||||||||
Db 51 tgctaagccagaccggtcctcg 72

RESULT 12
AAT66977
ID AAT66977 standard; DNA; 30 BP.
XX
AC AAT66977;
XX
DT 21-JUL-1997 (first entry)
XX
DE Cysteine protease CMH-1 PCR primer.
XX
KW Cysteine protease; CMH-1; CPP32/Mch2 homologue-1; apoptosis;
KW programmed cell death; cancer; neurodegenerative disease;
KW autoimmune disease; gene therapy; diagnosis; primer; PCR;
KW polymerase chain reaction; ss.
XX
OS Synthetic.
XX
PN WO9716552-A1.
XX
PD 09-MAY-1997.
XX
PF 01-NOV-1996; 96WO-US17431.
XX
PR 16-NOV-1995; 95US-0558733.
PR 03-NOV-1995; 95US-0007211.
PR 06-NOV-1995; 95US-0007251.
XX
PA (VERT-) VERTEX PHARM INC.
XX
PI Lippke JA, Su M;
XX
DR WPI; 1997-272121/24.
XX
PT DNA encoding active, activatable or inactive cysteine protease CMH-1
PT - useful in gene therapy for promoting and inhibiting apoptosis and
PT for diagnosing cells with potential for apoptosis
XX
PS Example 3; Page 25; 48pp; English.
XX
CC PCR primers (AAT66977-78) were designed to introduce in-frame XhoI
CC sites to the 5' and 3' ends of human cysteine protease cmh-1
CC cDNA (see also AAT66970). The PCR product XhoI fragment was ligated
CC into pEt-15b inducible E. coli expression vector. The resulting
CC plasmid encoded a polypeptide comprising a 23-residue N-terminal
CC peptide (AAW15248) contg. a hexahistidine tag and thrombin cleavage
CC site, fused to Ala24 of CHM-1. Hexahistidine-tagged CMH-1 (Ala24-
CC Gln303) was produced in E. coli BM21 (DE3) transformants.
XX
SQ Sequence 30 BP; 5 A; 11 C; 10 G; 4 T; 0 other;

Query Match 1.5%; Score 21; DB 18; Length 30;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 409 GCTAAGCCAGACCGGTCCTCG 429
|||||||||||||||||||||
Db 10 gctaagccagaccggtcctcg 30

RESULT 13
AAT15278
ID AAT15278 standard; cDNA; 31 BP.
XX
AC AAT15278;
XX
DT 28-OCT-1996 (first entry)
XX
DE Human ICE-LAP-3 5' oligonucleotide DNA primer.
XX
KW ICE-LAP-3; interleukin-I-converting enzyme-like apoptosis protease;
KW enzyme; Alzheimer's disease; Parkinson's disease; septic shock;
KW head injury; rheumatoid arthritis; PCR; polymerae chain reaction;
KW DNA primer; ss.
XX
OS Synthetic.
XX
PN WO9613603-A1.
XX
PD 09-MAY-1996.
XX
PF 06-JUN-1995; 95WO-US07235.
XX
PR 01-NOV-1994; 94US-0334251.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Hastings GA, He W, Hudson PL, Rosen CA;
XX
DR WPI; 1996-239509/24.
XX
PT Human ICE-LAP-3 and -4 DNA and protein - useful in the diagnosis
PT and treatment of Alzheimer's disease, Parkinson's disease,
PT rheumatoid arthritis, septic shock and head injury
XX
PS Disclosure; Page 30; 67pp; English.
XX
CC This 5' oligonucleotide DNA primer was used for the amplification
CC of the ICE-LAP-3 (ATCC 75875) DNA sequence, in conjunction with a
CC 3' DNA primer (AAT15279). This 5' DNA primer contains a BamHI site
CC followed by 18 nucleotides of ICE-LAP-3 coding sequence starting
CC from the presumed terminal AA of the processed protein codon.
CC The restriction site corresponds to the restriction site on
CC bacterial expression vector pQE-9.
XX
SQ Sequence 31 BP; 5 A; 8 C; 13 G; 5 T; 0 other;

Query Match 1.5%; Score 21; DB 17; Length 31;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 115 ATGCGTGCGGGGACACGGGTC 135

```
          ||||||||||||||||||||
Db     11 atgcgtgcggggacacgggtc 31


RESULT  14
AAT15279/c
ID   AAT15279 standard; cDNA; 31 BP.
XX
AC   AAT15279;
XX
DT   28-OCT-1996  (first entry)
XX
DE   Human ICE-LAP-3 3' oligonucleotide DNA primer.
XX
KW   ICE-LAP-3; interleukin-I-converting enzyme-like apoptosis protease;
KW   enzyme; Alzheimer's disease; Parkinson's disease; septic shock;
KW   head injury; rheumatoid arthritis; PCR; polymerae chain reaction;
KW   DNA primer; ss.
XX
OS   Synthetic.
XX
PN   WO9613603-A1.
XX
PD   09-MAY-1996.
XX
PF   06-JUN-1995;   95WO-US07235.
XX
PR   01-NOV-1994;   94US-0334251.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Hastings GA,  He W,  Hudson PL,  Rosen CA;
XX
DR   WPI; 1996-239509/24.
XX
PT   Human ICE-LAP-3 and -4 DNA and protein  - useful in the diagnosis
PT   and treatment of Alzheimer's disease, Parkinson's disease,
PT   rheumatoid arthritis, septic shock and head injury
XX
PS   Disclosure; Page 30; 67pp; English.
XX
CC   This 3' oligonucleotide DNA primer was used for the amplification
CC   of the ICE-LAP-3 (ATCC 75875) DNA sequence, in conjunction with a
CC   5' DNA primer (AAT15278).  This 3' DNA primer contains complementary
CC   sequences to an XbaI site followed by 21 nuclotides of ICE-LAP-3.
CC   The restriction site corresponds to the site on the bacterial
CC   expression vector pQE-9.
XX
SQ   Sequence 31 BP; 7 A; 7 C; 8 G; 9 T; 0 other;

  Query Match             1.5%;  Score 21;  DB 17;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 6;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1120 ATCCTCCACCAGGGTGAATGA 1140
           |||||||||||||||||||||
Db     31 ATCCTCCACCAGGGTGAATGA 11


RESULT  15
AAT66974/c
ID   AAT66974 standard; DNA; 33 BP.
XX
AC   AAT66974;
XX
DT   21-JUL-1997  (first entry)
XX
DE   T7-tagged CMH-1 PCR primer.
XX
KW   Cysteine protease; CMH-1; CPP32/Mch2 homologue-1; apoptosis;
KW   programmed cell death; cancer; neurodegenerative disease;
KW   autoimmune disease; gene therapy; diagnosis; primer; PCR;
KW   polymerase chain reaction; ss.
XX
OS   Synthetic.
XX
PN   WO9716552-A1.
XX
PD   09-MAY-1997.
XX
PF   01-NOV-1996;   96WO-US17431.
XX
PR   16-NOV-1995;   95US-0558733.
PR   03-NOV-1995;   95US-0007211.
PR   06-NOV-1995;   95US-0007251.
XX
PA   (VERT-) VERTEX PHARM INC.
XX
PI   Lippke JA,  Su M;
XX
DR   WPI; 1997-272121/24.
XX
PT   DNA encoding active, activatable or inactive cysteine protease CMH-1
PT   - useful in gene therapy for promoting and inhibiting apoptosis and
PT   for diagnosing cells with potential for apoptosis
XX
PS   Example 2; Page 23; 48pp; English.
XX
CC   PCR primers (AAT66973-74) were used to produce a cDNA fragment
CC   encoding an N-terminally T7-tagged cysteine protein CMH-1 lacking
CC   the first 23 amino acid residues of the proenzyme (see also
CC   AAW15247).  Amplified DNA was subcloned into an SV40-based expression
CC   vector under control of the SR-alpha promoter.  COS cells
CC   transiently transfected with the vector resulted in high
CC   expression of a truncated CMH-1 protein.  Over-expression of
CC   the protein resulted in cell apoptosis.
XX
SQ   Sequence 33 BP; 10 A; 4 C; 7 G; 12 T; 0 other;

  Query Match             1.5%;  Score 21;  DB 18;  Length 33;
  Best Local Similarity   100.0%;  Pred. No. 6;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1232 GAACTCTACTTCAGTCAATAG 1252
           |||||||||||||||||||||
Db     33 GAACTCTACTTCAGTCAATAG 13


Search completed: August 16, 2002, 01:48:42
Job time: 6183 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2002, 23:59:49 ; Search time 75.09 Seconds
(without alignments)
4484.799 Million cell updates/sec

Title: US-09-613-508-1
Perfect score: 1371
Sequence: 1 GCACGAGAAACTTTGCTGTG.....TAATTTCAGGAAAATGTATT 1371

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1371	100.0	1371	3	US-08-462-969B-1	Sequence 1, Appli
c	2	28	2.0	28	2	US-08-859-998-1322	Sequence 1322, Ap
c	3	28	2.0	28	4	US-09-225-928-1322	Sequence 1322, Ap
	4	26	1.9	26	2	US-08-859-998-1321	Sequence 1321, Ap
	5	26	1.9	26	4	US-09-225-928-1321	Sequence 1321, Ap
	6	22	1.6	22	3	US-08-462-969B-9	Sequence 9, Appli
c	7	22	1.6	53	3	US-08-462-969B-10	Sequence 10, Appl
	8	21	1.5	31	3	US-08-462-969B-5	Sequence 5, Appli
c	9	21	1.5	31	3	US-08-462-969B-6	Sequence 6, Appli
c	10	19	1.4	435	1	US-08-592-126-130	Sequence 130, App
c	11	19	1.4	4635	2	US-08-404-531B-27	Sequence 27, Appl
c	12	19	1.4	4635	3	US-08-476-900A-27	Sequence 27, Appl
c	13	19	1.4	4635	3	US-08-488-546A-27	Sequence 27, Appl
c	14	19	1.4	5110	2	US-08-404-531B-4	Sequence 4, Appli
c	15	19	1.4	5110	2	US-08-404-531B-5	Sequence 5, Appli
c	16	19	1.4	5110	3	US-08-476-900A-4	Sequence 4, Appli
c	17	19	1.4	5110	3	US-08-476-900A-5	Sequence 5, Appli
c	18	19	1.4	5110	3	US-08-488-546A-4	Sequence 4, Appli
c	19	19	1.4	5110	3	US-08-488-546A-5	Sequence 5, Appli
	20	18	1.3	3369	2	US-08-619-198-2	Sequence 2, Appli
	21	18	1.3	4170	2	US-08-619-198-4	Sequence 4, Appli
	22	18	1.3	5192	2	US-08-619-198-8	Sequence 8, Appli
	23	18	1.3	5198	2	US-08-619-198-1	Sequence 1, Appli
	24	17	1.2	1531	1	US-08-449-986-1	Sequence 1, Appli
	25	17	1.2	1531	2	US-08-756-855-1	Sequence 1, Appli
c	26	17	1.2	1768	3	US-09-150-133-2	Sequence 2, Appli
c	27	17	1.2	1768	3	US-09-150-141-2	Sequence 2, Appli
c	28	17	1.2	1768	4	US-09-374-493-2	Sequence 2, Appli
c	29	17	1.2	1768	4	US-09-374-824-2	Sequence 2, Appli
c	30	17	1.2	1768	4	US-09-374-492-2	Sequence 2, Appli
	31	17	1.2	2686	4	US-09-228-986-3	Sequence 3, Appli
c	32	17	1.2	2815	1	US-08-230-491A-1	Sequence 1, Appli
c	33	17	1.2	2815	1	US-08-619-280A-1	Sequence 1, Appli
c	34	17	1.2	2815	2	US-08-940-391-1	Sequence 1, Appli
c	35	17	1.2	3061	4	US-09-147-119-6	Sequence 6, Appli
c	36	17	1.2	3116	2	US-08-449-645A-14	Sequence 14, Appl
c	37	17	1.2	3116	2	US-08-702-367A-14	Sequence 14, Appl
c	38	17	1.2	3116	5	PCT-US95-04681-14	Sequence 14, Appl
c	39	17	1.2	3348	1	US-08-222-616-34	Sequence 34, Appl
c	40	17	1.2	3348	5	PCT-US95-04228-34	Sequence 34, Appl
	41	17	1.2	3414	1	US-08-764-100-17	Sequence 17, Appl
c	42	17	1.2	3414	1	US-08-764-100-21	Sequence 21, Appl
	43	17	1.2	4866	1	US-08-110-158-5	Sequence 5, Appli
	44	17	1.2	4970	1	US-08-764-100-14	Sequence 14, Appl
c	45	17	1.2	4970	1	US-08-764-100-20	Sequence 20, Appl

ALIGNMENTS

```
RESULT   1
US-08-462-969B-1
; Sequence 1, Application US/08462969B
; Patent No. 6087150
;  GENERAL INFORMATION:
;    APPLICANT:  He, Wei-Wu et al.
;    TITLE OF INVENTION:  Interleukin-1 Beta Converting Enzyme
;    TITLE OF INVENTION:  Like Apoptosis Protease 3 and 4
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Ave.
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,969B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/334,251
;      FILING DATE:  11-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF140P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1371 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-462-969B-1


  Query Match           100.0%;  Score 1371;  DB 3;  Length 1371;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1371;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCACGAGAAACTTTGCTGTGCGCGTTCTCCCGCGCGCGGGCTCAACTTTGTAGAGCGAGG 60
```

```
Db      1 GCACGAGAAACTTTGCTGTGCGCGTTCTCCCGCGCGCGGGCTCAACTTTGTAGAGCGAGG 60

Qy     61 GGCCAACTTGGCAGAGCGCGCGGCCAGCTTTGCAGAGAGCGCCCTCCAGGGACTATGCGT 120

Db     61 GGCCAACTTGGCAGAGCGCGCGGCCAGCTTTGCAGAGAGCGCCCTCCAGGGACTATGCGT 120

Qy    121 GCGGGGACACGGGTCGCTTTGGGCTCTTCCACCCCTGCGGAGCGCACTACCCCGAGCCAG 180

Db    121 GCGGGGACACGGGTCGCTTTGGGCTCTTCCACCCCTGCGGAGCGCACTACCCCGAGCCAG 180

Qy    181 GGGCGGTGCAAGCCCCGCCCGGCCCTACCCAGGGCGGCTCCTCCCTCCGCAGCGCCGAGA 240

Db    181 GGGCGGTGCAAGCCCCGCCCGGCCCTACCCAGGGCGGCTCCTCCCTCCGCAGCGCCGAGA 240

Qy    241 CTTTTAGTTTCGCTTTCGCTAAAGGGGCCCCAGACCCTTGCTGCGGAGCGACGGAGAGAG 300

Db    241 CTTTTAGTTTCGCTTTCGCTAAAGGGGCCCCAGACCCTTGCTGCGGAGCGACGGAGAGAG 300

Qy    301 ACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGGGCTGT 360

Db    301 ACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGGGCTGT 360

Qy    361 ATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGAC 420

Db    361 ATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGAC 420

Qy    421 CGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCC 480

Db    421 CGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCC 480

Qy    481 ATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTG 540

Db    481 ATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTG 540

Qy    541 GGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGA 600

Db    541 GGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGA 600

Qy    601 AACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGAC 660

Db    601 AACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGAC 660

Qy    661 GTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCT 720

Db    661 GTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCT 720

Qy    721 GAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAA 780

Db    721 GAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAA 780

Qy    781 AATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGG 840

Db    781 AATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGG 840

Qy    841 GGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGA 900

Db    841 GGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGA 900

Qy    901 GGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCT 960

Db    901 GGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCT 960

Qy    961 AATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCA 1020

Db    961 AATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCA 1020

Qy   1021 GGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCC 1080

Db   1021 GGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCC 1080

Qy   1081 ATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGA 1140

Db   1081 ATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGA 1140

Qy   1141 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCA 1200

Db   1141 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCA 1200

Qy   1201 GATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCA 1260

Db   1201 GATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCA 1260

Qy   1261 GGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGC 1320

Db   1261 GGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGC 1320

Qy   1321 TCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371

Db   1321 TCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
```

RESULT 2
US-08-859-998-1322/c
; Sequence 1322, Application US/08859998
; Patent No. 5994076
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Jokhadze, George
; APPLICANT: Bibilashvilli, Robert
; TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1375
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: CA
; COUNTRY: US
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/859,998
; FILING DATE: 21-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Field, Bret E.
; REGISTRATION NUMBER: 37,620
; REFERENCE/DOCKET NUMBER: 09096/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 1322:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; OTHER INFORMATION: oligonucleotide primer
US-08-859-998-1322

Query Match 2.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 6.6e-05;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   1156 CTTTGAGTCTCAGTCTGATGACCCACAC 1183
```

```
          ||||||||||||||||||||||||||||
Db      28 CTTTGAGTCTCAGTCTGATGACCCACAC 1


RESULT   3
US-09-225-928-1322/c
; Sequence 1322, Application US/09225928
; Patent No. 6352829
;   GENERAL INFORMATION:
;        APPLICANT: Chenchik, Alex
;                   Jokhadze, George
;                   Bibilashvilli, Robert
;        TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
;                            EXPRESSION
;        NUMBER OF SEQUENCES: 1375
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson, P.C.
;             STREET: 2200 Sand Hill Road, Suite 100
;             CITY: Menlo Park
;             STATE: CA
;             COUNTRY: US
;             ZIP: 94025
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows95
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/225,928
;             FILING DATE: 05-Jan-1999
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/859,998
;             FILING DATE: 21-MAY-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Field, Bret E.
;             REGISTRATION NUMBER: 37,620
;             REFERENCE/DOCKET NUMBER: 09096/002001
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415-322-5070
;             TELEFAX: 415-854-0875
;   INFORMATION FOR SEQ ID NO: 1322:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 28 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA
;        FEATURE:
;             OTHER INFORMATION:  oligonucleotide primer
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1322:
US-09-225-928-1322


  Query Match             2.0%;  Score 28;  DB 4;  Length 28;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-05;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1156 CTTTGAGTCTCAGTCTGATGACCCACAC 1183
          ||||||||||||||||||||||||||||
Db     28 CTTTGAGTCTCAGTCTGATGACCCACAC 1


RESULT   4
US-08-859-998-1321
; Sequence 1321, Application US/08859998
; Patent No. 5994076
;  GENERAL INFORMATION:
;    APPLICANT:  Chenchik, Alex
;    APPLICANT:  Jokhadze, George
;    APPLICANT:  Bibilashvilli, Robert
;    TITLE OF INVENTION:  METHOD OF ASSAYING DIFFERENTIAL
;    TITLE OF INVENTION:  EXPRESSION
;    NUMBER OF SEQUENCES:  1375
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  2200 Sand Hill Road, Suite 100
;      CITY:  Menlo Park
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/859,998
;      FILING DATE:  21-MAY-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Field, Bret E.
;      REGISTRATION NUMBER:  37,620
;      REFERENCE/DOCKET NUMBER:  09096/002001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-322-5070
;      TELEFAX:  415-854-0875
;  INFORMATION FOR SEQ ID NO:  1321:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    FEATURE:
;      OTHER INFORMATION:  oligonucleotide primer
US-08-859-998-1321


  Query Match             1.9%;  Score 26;  DB 2;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 0.00071;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    880 TTCTTCATTCAGGCTTGCCGAGGGAC 905
          ||||||||||||||||||||||||||
Db      1 TTCTTCATTCAGGCTTGCCGAGGGAC 26


RESULT   5
US-09-225-928-1321
; Sequence 1321, Application US/09225928
; Patent No. 6352829
;   GENERAL INFORMATION:
;        APPLICANT: Chenchik, Alex
;                   Jokhadze, George
;                   Bibilashvilli, Robert
;        TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
;                            EXPRESSION
;        NUMBER OF SEQUENCES: 1375
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson, P.C.
;             STREET: 2200 Sand Hill Road, Suite 100
;             CITY: Menlo Park
;             STATE: CA
;             COUNTRY: US
;             ZIP: 94025
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows95
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/225,928
```

```
;             FILING DATE: 05-Jan-1999
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/859,998
;             FILING DATE: 21-MAY-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Field, Bret E.
;             REGISTRATION NUMBER: 37,620
;             REFERENCE/DOCKET NUMBER: 09096/002001
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415-322-5070
;             TELEFAX: 415-854-0875
;   INFORMATION FOR SEQ ID NO: 1321:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 26 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA
;        FEATURE:
;             OTHER INFORMATION:  oligonucleotide primer
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1321:
US-09-225-928-1321


  Query Match             1.9%;  Score 26;  DB 4;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 0.00071;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    880 TTCTTCATTCAGGCTTGCCGAGGGAC 905
          ||||||||||||||||||||||||||
Db      1 TTCTTCATTCAGGCTTGCCGAGGGAC 26


RESULT   6
US-08-462-969B-9
; Sequence 9, Application US/08462969B
; Patent No. 6087150
;  GENERAL INFORMATION:
;    APPLICANT:  He, Wei-Wu et al.
;    TITLE OF INVENTION:  Interleukin-1 Beta Converting Enzyme
;    TITLE OF INVENTION:  Like Apoptosis Protease 3 and 4
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Ave.
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,969B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/334,251
;      FILING DATE:  11-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF140P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-462-969B-9


  Query Match             1.6%;  Score 22;  DB 3;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.083;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    111 GACTATGCGTGCGGGGACACGG 132
          ||||||||||||||||||||||
Db      1 GACTATGCGTGCGGGGACACGG 22


RESULT   7
US-08-462-969B-10/c
; Sequence 10, Application US/08462969B
; Patent No. 6087150
;  GENERAL INFORMATION:
;    APPLICANT:  He, Wei-Wu et al.
;    TITLE OF INVENTION:  Interleukin-1 Beta Converting Enzyme
;    TITLE OF INVENTION:  Like Apoptosis Protease 3 and 4
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Ave.
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,969B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/334,251
;      FILING DATE:  11-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF140P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  53 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-462-969B-10


  Query Match             1.6%;  Score 22;  DB 3;  Length 53;
  Best Local Similarity   100.0%;  Pred. No. 0.087;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1117 CAAATCCTCCACCAGGGTGAAT 1138
          ||||||||||||||||||||||
Db     53 CAAATCCTCCACCAGGGTGAAT 32


RESULT   8
US-08-462-969B-5
; Sequence 5, Application US/08462969B
; Patent No. 6087150
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  He, Wei-Wu et al.
;    TITLE OF INVENTION:  Interleukin-1 Beta Converting Enzyme
;    TITLE OF INVENTION:  Like Apoptosis Protease 3 and 4
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Ave.
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,969B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/334,251
;      FILING DATE:  11-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF140P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-462-969B-5


  Query Match             1.5%;  Score 21;  DB 3;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 0.28;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    115 ATGCGTGCGGGGACACGGGTC 135
          |||||||||||||||||||||
Db     11 ATGCGTGCGGGGACACGGGTC 31


RESULT   9
US-08-462-969B-6/c
; Sequence 6, Application US/08462969B
; Patent No. 6087150
;  GENERAL INFORMATION:
;    APPLICANT:  He, Wei-Wu et al.
;    TITLE OF INVENTION:  Interleukin-1 Beta Converting Enzyme
;    TITLE OF INVENTION:  Like Apoptosis Protease 3 and 4
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Ave.
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,969B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/334,251
;      FILING DATE:  11-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF140P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-462-969B-6


  Query Match             1.5%;  Score 21;  DB 3;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 0.28;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1120 ATCCTCCACCAGGGTGAATGA 1140
          |||||||||||||||||||||
Db     31 ATCCTCCACCAGGGTGAATGA 11


RESULT  10
US-08-592-126-130/c
; Sequence 130, Application US/08592126
; Patent No. 5821091
;  GENERAL INFORMATION:
;    APPLICANT:  Gregory Dolganov
;    TITLE OF INVENTION:  Transcripts Encoding Immunomodulatory
;    TITLE OF INVENTION:  Polypeptides
;    NUMBER OF SEQUENCES:  151
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  350 Cambridge Avenue, Suite 250
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,126
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sholtz, Charles K.
;      REGISTRATION NUMBER:  38,615
;      REFERENCE/DOCKET NUMBER:  4600-0111
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  130:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  435 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
```

```
;      INDIVIDUAL ISOLATE:  G75.seq
US-08-592-126-130


  Query Match             1.4%;  Score 19;  DB 1;  Length 435;
  Best Local Similarity   100.0%;  Pred. No. 3.5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1097 GAAAAGACCTGGAAATCAT 1115
          |||||||||||||||||||
Db    208 GAAAAGACCTGGAAATCAT 190


RESULT  11
US-08-404-531B-27/c
; Sequence 27, Application US/08404531B
; Patent No. 5863724
;  GENERAL INFORMATION:
;    APPLICANT:  Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
;    APPLICANT:  Thomas, Gilbert Cote, and Robert Gagel
;    TITLE OF INVENTION:  Sequence Encoding Mammalian Sulfonylurea Receptor          a
; Patent No. 5863724
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz &
;      ADDRESSEE:  No. 5863724ris
;      STREET:  One Liberty Place 46th. Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/404,531B
;      FILING DATE:  15-MAR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Beardell, Lori Y.
;      REGISTRATION NUMBER:  34,293
;      REFERENCE/DOCKET NUMBER:  BYLR-0003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4635 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  37..4533
US-08-404-531B-27


  Query Match             1.4%;  Score 19;  DB 2;  Length 4635;
  Best Local Similarity   100.0%;  Pred. No. 4;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    912 TGATGATGGCATCCAGGCC 930
          |||||||||||||||||||
Db   4210 TGATGATGGCATCCAGGCC 4192


RESULT  12
US-08-476-900A-27/c
; Sequence 27, Application US/08476900A
; Patent No. 6031150
;  GENERAL INFORMATION:
;    APPLICANT:  Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson
;    TITLE OF INVENTION:  Sequence Encoding Mammalian Sulfonylurea Receptor
; Patent No. 6031150
;    TITLE OF INVENTION:  and Method of Detecting Persistent Hyperinsulinemic Hypogly
;    TITLE OF INVENTION:  Infancy
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz &
;      ADDRESSEE:  No. 6031150ris
;      STREET:  One Liberty Place 46th. Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,900A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Beardell, Lori Y.
;      REGISTRATION NUMBER:  34,293
;      REFERENCE/DOCKET NUMBER:  BYLR-0027
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4635 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  37..4533
US-08-476-900A-27


  Query Match             1.4%;  Score 19;  DB 3;  Length 4635;
  Best Local Similarity   100.0%;  Pred. No. 4;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    912 TGATGATGGCATCCAGGCC 930
          |||||||||||||||||||
Db   4210 TGATGATGGCATCCAGGCC 4192


RESULT  13
US-08-488-546A-27/c
; Sequence 27, Application US/08488546A
; Patent No. 6054313
;  GENERAL INFORMATION:
;    APPLICANT:  Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
;    APPLICANT:  Thomas, Gilbert Cote, and Robert Gagel
;    TITLE OF INVENTION:  Sequence Encoding Mammalian Sulfonylurea Receptor
; Patent No. 6054313
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz &
;      ADDRESSEE:  No. 6054313ris
;      STREET:  One Liberty Place 46th. Floor
```

```
;       CITY:  Philadelphia
;       STATE:  PA
;       COUNTRY:  USA
;       ZIP:  19103
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/488,546A
;       FILING DATE:  07-JUNE-1995
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/404,531
;       FILING DATE:  15-MARCH-1995
;       CLASSIFICATION:  800
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Beardell, Lori Y.
;       REGISTRATION NUMBER:  34,293
;       REFERENCE/DOCKET NUMBER:  BYLR-0026
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  215-568-3100
;       TELEFAX:  215-568-3439
;   INFORMATION FOR SEQ ID NO:  27:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4635 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     FRAGMENT TYPE:  N-terminal
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  37..4533
US-08-488-546A-27


  Query Match             1.4%;  Score 19;  DB 3;  Length 4635;
  Best Local Similarity   100.0%;  Pred. No. 4;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     912 TGATGATGGCATCCAGGCC 930
           ||||||||||||||||||| 
Db    4210 TGATGATGGCATCCAGGCC 4192


RESULT  14
US-08-404-531B-4/c
; Sequence 4, Application US/08404531B
; Patent No. 5863724
;  GENERAL INFORMATION:
;    APPLICANT:  Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
;    APPLICANT:  Thomas, Gilbert Cote, and Robert Gagel
;    TITLE OF INVENTION:  Sequence Encoding Mammalian Sulfonylurea Receptor          a
; Patent No. 5863724
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz &
;      ADDRESSEE:  No. 5863724ris
;      STREET:  One Liberty Place 46th. Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/404,531B
;      FILING DATE:  15-MAR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Beardell, Lori Y.
;      REGISTRATION NUMBER:  34,293
;      REFERENCE/DOCKET NUMBER:  BYLR-0003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5110 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  nucleic acid
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-404-531B-4


  Query Match             1.4%;  Score 19;  DB 2;  Length 5110;
  Best Local Similarity   100.0%;  Pred. No. 4.1;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     912 TGATGATGGCATCCAGGCC 930
           ||||||||||||||||||| 
Db    4685 TGATGATGGCATCCAGGCC 4667


RESULT  15
US-08-404-531B-5/c
; Sequence 5, Application US/08404531B
; Patent No. 5863724
;  GENERAL INFORMATION:
;    APPLICANT:  Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
;    APPLICANT:  Thomas, Gilbert Cote, and Robert Gagel
;    TITLE OF INVENTION:  Sequence Encoding Mammalian Sulfonylurea Receptor
; Patent No. 5863724
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz &
;      ADDRESSEE:  No. 5863724ris
;      STREET:  One Liberty Place 46th. Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/404,531B
;      FILING DATE:  15-MAR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Beardell, Lori Y.
;      REGISTRATION NUMBER:  34,293
;      REFERENCE/DOCKET NUMBER:  BYLR-0003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5110 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
```

```
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  260..5004
US-08-404-531B-5


  Query Match             1.4%;  Score 19;  DB 2;  Length 5110;
  Best Local Similarity   100.0%;  Pred. No. 4.1;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     912 TGATGATGGCATCCAGGCC 930
           ||||||||||||||||||
Db    4685 TGATGATGGCATCCAGGCC 4667


Search completed: August 16, 2002, 01:43:01
Job time: 6192 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2002, 23:02:24 ; Search time 2424.61 Seconds
(without alignments)
7631.877 Million cell updates/sec

Title: US-09-613-508-1
Perfect score: 1371
Sequence: 1 GCACGAGAAACTTTGCTGTG.....TAATTTCAGGAAAATGTATT 1371

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	768	56.0	973	9	AL542189	AL542189 AL542189
2	641	46.8	745	10	BI824214	BI824214 603040584
3	589	43.0	830	10	BG025549	BG025549 602274928
4	540	39.4	774	9	AU130224	AU130224 AU130224
5	533	38.9	656	10	BF971443	BF971443 602272969
6	508	37.1	926	10	BI821936	BI821936 603039745
7	499	36.4	809	9	AU126220	AU126220 AU126220
8	489	35.7	719	9	AU143139	AU143139 AU143139
9	465	33.9	791	9	AU134778	AU134778 AU134778
10	424	30.9	1006	10	BG339524	BG339524 602437368
11	416	30.3	796	10	BE619147	BE619147 601472883
12	415	30.3	497	10	BM194512	BM194512 TCAAP2Q11
13	413	30.1	700	9	AU117167	AU117167 AU117167
14	392	28.6	485	10	BM147460	BM147460 TCAAP1Q11
15	392	28.6	532	10	BM150092	BM150092 TCAAP4Q11
16	386	28.2	621	10	BE894817	BE894817 601436392
17	384	28.0	1075	10	BM459840	BM459840 AGENCOURT
18	364	26.5	961	10	BE891855	BE891855 601435409
19	360	26.3	630	10	BE893968	BE893968 601437754
20	334	24.4	808	10	BF540869	BF540869 602067587
21	334	24.4	840	9	AU131461	AU131461 AU131461
22	313	22.8	917	10	BG674973	BG674973 602621183
23	253	18.5	402	9	AA101428	AA101428 zn72e08.r
24	239	17.4	290	9	AA297423	AA297423 EST112974
25	231	16.8	334	9	AA298931	AA298931 EST114527
26	222	16.2	356	9	AA298267	AA298267 EST114122
27	187	13.6	846	10	BF668893	BF668893 602123296
c 28	174	12.7	486	9	AI909733	AI909733 IL-BT223-
29	152	11.1	593	10	H91868	H91868 ys81a06.r1
30	150	10.9	756	10	BI458770	BI458770 603200030
31	126	9.2	239	9	AA297274	AA297274 EST112969
32	124	9.0	226	9	AA296944	AA296944 EST112526
33	120	8.8	923	10	BF698196	BF698196 602125529
34	112	8.2	244	10	BE711179	BE711179 RC6-HT067
35	112	8.2	334	10	BE775078	BE775078 IL2-UM007
36	112	8.2	334	10	BE775099	BE775099 IL2-UM007
37	101	7.4	277	10	T50828	T50828 yb31c06.r1
38	72	5.3	588	10	BG248784	BG248784 602400969
c 39	44	3.2	154	10	BG384410	BG384410 303485 MA
40	32	2.3	223	9	AW347772	AW347772 31988 MAR
41	32	2.3	360	10	BG967610	BG967610 602833403
42	32	2.3	394	10	BG918158	BG918158 602821076
43	32	2.3	412	9	AA939788	AA939788 vz92g09.r
44	32	2.3	473	9	BB862907	BB862907 BB862907
45	32	2.3	588	10	BI401667	BI401667 MI-P-CP0-

ALIGNMENTS

RESULT 1
AL542189
LOCUS AL542189 973 bp mRNA linear EST 16-FEB-2001
DEFINITION AL542189 LTI_FL002_PL1 Homo sapiens cDNA clone CS0DE008YG21 5 prime , mRNA sequence.
ACCESSION AL542189
VERSION AL542189.1 GI:12873988
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 973)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES Location/Qualifiers
source 1. .973
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DE008YG21"
/clone_lib="LTI_FL002_PL1"
/lab_host="DH10B"
/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville , Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 275 a 239 c 241 g 217 t 1 others
ORIGIN

Query Match 56.0%; Score 768; DB 9; Length 973;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 768; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 105

Qy    411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 165

Qy    471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    166 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 225

Qy    531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    226 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 285

Qy    591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    286 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 345

Qy    651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    346 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 405

Qy    711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    406 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 465

Qy    771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    466 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 525

Qy    831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    526 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 585

Qy    891 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 950
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    586 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 645

Qy    951 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 1010
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    646 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 705

Qy   1011 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 1070
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    706 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 765

Qy   1071 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCA 1118
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    766 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCA 813
```

RESULT 2
BI824214
LOCUS BI824214 745 bp mRNA linear EST 04-OCT-2001
DEFINITION 603040584F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5181381 5', mRNA sequence.
ACCESSION BI824214
VERSION BI824214.1 GI:15935764
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 745)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11452 row: h column: 22
High quality sequence stop: 740.
FEATURES Location/Qualifiers
source 1. .745
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5181381"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."
BASE COUNT 215 a 178 c 185 g 167 t
ORIGIN

Query Match 46.8%; Score 641; DB 10; Length 745;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 691; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy    351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     53 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 112

Qy    411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    113 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 172

Qy    471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    173 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 232

Qy    531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    233 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 292

Qy    591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    293 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 352

Qy    651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    353 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 412

Qy    711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    413 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 472

Qy    771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    473 TGGAGAAGAACATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 532

Qy    831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    533 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 592
```

```
Qy    891 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 950
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    593 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 652

Qy    951 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 1010
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    653 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 712

Qy   1011 CACGGTTCCAGGCTATTACTCGTGGAGGAGCC 1042
          ||||||||||||||||||||||||||||||||
Db    713 CACGGTTCCAGGCTATTACTCGTGGAGGAGCC 744


RESULT   3
BG025549
LOCUS       BG025549                 830 bp    mRNA    linear   EST 24-JAN-2001
DEFINITION  602274928F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4363038 5',
            mRNA sequence.
ACCESSION   BG025549
VERSION     BG025549.1  GI:12412258
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 830)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis Staudt, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10008  row: g  column: 07
            High quality sequence stop: 669.
FEATURES             Location/Qualifiers
     source          1. .830
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4363038"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      229 a    206 c    196 g    197 t      2 others
ORIGIN


  Query Match             43.0%;  Score 589;  DB 10;  Length 830;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  589;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    502 GTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAAC 561
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAAC 60

Qy    562 AACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCC 621
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCC 120

Qy    622 GAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGC 681
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGC 180

Qy    682 TCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCC 741
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCC 240

Qy    742 GCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGAT 801
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGAT 300

Qy    802 GGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTT 861
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTT 360

Qy    862 TTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGC 921
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGC 420

Qy    922 ATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCA 981
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCA 480

Qy    982 GTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGC 1041
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGC 540

Qy   1042 CCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGG 1090
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGG 589


RESULT   4
AU130224
LOCUS       AU130224                 774 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  AU130224 NT2RP3 Homo sapiens cDNA clone NT2RP3000452 5', mRNA
            sequence.
ACCESSION   AU130224
VERSION     AU130224.1  GI:10990578
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 774)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .774
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP3000452"
                     /clone_lib="NT2RP3"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                     cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT      220 a    179 c    193 g    178 t      4 others
ORIGIN


  Query Match             39.4%;  Score 540;  DB 9;  Length 774;
```

```
  Best Local Similarity   100.0%;  Pred. No. 1.6e-289;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 173

Qy     411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     174 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 233

Qy     471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     234 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 293

Qy     531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     294 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 353

Qy     591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     354 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 413

Qy     651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     414 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 473

Qy     711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     474 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 533

Qy     771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     534 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 593

Qy     831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     594 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 653


RESULT   5
BF971443
LOCUS       BF971443                 656 bp    mRNA    linear   EST 22-JAN-2001
DEFINITION  602272969F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4361176 5',
            mRNA sequence.
ACCESSION   BF971443
VERSION     BF971443.1  GI:12338658
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 656)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10003  row: i  column: 17
            High quality sequence stop: 649.
FEATURES             Location/Qualifiers
     source          1. .656
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4361176"
                     /clone_lib="NIH_MGC_84"
                     /tissue_type="adrenal cortex carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site_1:
                     NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT
                     primed. Average insert size 1.229 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      188 a    155 c    161 g    152 t
ORIGIN


  Query Match             38.9%;  Score 533;  DB 10;  Length 656;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-285;
  Matches  533;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     525 GAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGAC 584
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGAC 60

Qy     585 AGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCG 644
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCG 120

Qy     645 AAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCT 704
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCT 180

Qy     705 GCTTAAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTT 764
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GCTTAAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTT 240

Qy     765 AAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTT 824
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTT 300

Qy     825 GACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTT 884
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTT 360

Qy     885 CATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCAT 944
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCAT 420

Qy     945 CAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGC 1004
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGC 480

Qy    1005 CTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCT 1057
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCT 533


RESULT   6
BI821936
LOCUS       BI821936                 926 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603039745F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5180651 5',
            mRNA sequence.
ACCESSION   BI821936
VERSION     BI821936.1  GI:15933486
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 926)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

```
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM11450  row: j  column: 12
           High quality sequence stop: 882.
FEATURES             Location/Qualifiers
     source          1. .926
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5180651"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69.  Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon  cloning).  Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was  constructed by C.
                     Gruber (Invitrogen).  Research Genetics tracking  code
                     021. Note: this is a NIH_MGC Library."
BASE COUNT      264 a    218 c    242 g    202 t
ORIGIN


  Query Match             37.1%;  Score 508;  DB 10;  Length 926;
  Best Local Similarity   99.8%;  Pred. No. 1.2e-271;
  Matches  628;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy    352 CAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCT 411
Db     65 CAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCT 124

Qy    412 AAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACC 471
Db    125 AAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACC 184

Qy    472 ATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTAC-ATATCAGTACAACATGAATTT 530
Db    185 ATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACGATATCAGTACAACATGAATTT 244

Qy    531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
Db    245 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 304

Qy    591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
Db    305 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 364

Qy    651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
Db    365 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 424

Qy    711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
Db    425 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 484

Qy    771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
Db    485 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 544

Qy    831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
Db    545 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 604

Qy    891 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 950
Db    605 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 664

Qy    951 CACAGATGCTAATCCTCGATACAAGATCC 979
Db    665 CACAGATGCTAATCCTCGATACAAGATCC 693


RESULT   7
AU126220
LOCUS       AU126220                809 bp    mRNA    linear   EST 23-OCT-2000
DEFINITION  AU126220 NT2RP1 Homo sapiens cDNA clone NT2RP1000352 5', mRNA
            sequence.
ACCESSION   AU126220
VERSION     AU126220.1  GI:10950936
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 809)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Kawai,Y., Ishii,S., Saito,K.,
            Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
  TITLE     HRI human cDNA project (Ota,T., Nishikawa,T., Suzuki,Y., Kawai,Y.,
            Ishii,S., Saito,K., Nakamura,Y., Nagai,T., Sugano,S., Isogai,T.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .809
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP1000352"
                     /clone_lib="NT2RP1"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pUC19FL3; mRNA from NT2 neuronal precursor
                     cells after 48-hours retinoic acid (RA) induction"
BASE COUNT      230 a    197 c    197 g    178 t      7 others
ORIGIN


  Query Match             36.4%;  Score 499;  DB 9;  Length 809;
  Best Local Similarity   99.8%;  Pred. No. 1.2e-266;
  Matches  549;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
Db    130 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 189

Qy    411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
Db    190 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 249

Qy    471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
Db    250 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 309

Qy    531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
Db    310 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 369

Qy    591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
Db    370 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 429

Qy    651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
Db    430 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 489
```

```
Qy    711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    490 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 549

Qy    771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    550 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAANGATTTGACAGC 609

Qy    831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    610 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 669

Qy    891 GGCTTGCCGA 900
          ||||||||||
Db    670 GGCTTGCCGA 679
```

```
RESULT   8
AU143139
LOCUS       AU143139                719 bp    mRNA    linear   EST 25-OCT-2000
DEFINITION  AU143139 Y79AA1 Homo sapiens cDNA clone Y79AA1001448 5', mRNA
            sequence.
ACCESSION   AU143139
VERSION     AU143139.1  GI:11004660
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 719)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .719
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Y79AA1001448"
                     /clone_lib="Y79AA1"
                     /cell_type="retinoblastoma"
                     /cell_line="Y79"
                     /note="Vector: pME18SFL3"
BASE COUNT      212 a    169 c    181 g    154 t      3 others
ORIGIN
```

```
  Query Match             35.7%;  Score 489;  DB 9;  Length 719;
  Best Local Similarity   99.8%;  Pred. No. 4.6e-261;
  Matches  539;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 140

Qy    411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 200

Qy    471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 260

Qy    531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 TGAAAAGCCGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 320

Qy    591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 380

Qy    651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    381 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 440

Qy    711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    441 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 500

Qy    771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    501 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 560

Qy    831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    561 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 620
```

```
RESULT   9
AU134778
LOCUS       AU134778                791 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  AU134778 PLACE1 Homo sapiens cDNA clone PLACE1000485 5', mRNA
            sequence.
ACCESSION   AU134778
VERSION     AU134778.1  GI:10995317
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 791)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .791
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="PLACE1000485"
                     /clone_lib="PLACE1"
                     /tissue_type="placenta"
                     /note="Vector: pME18SFL3"
BASE COUNT      234 a    189 c    195 g    166 t      7 others
ORIGIN
```

```
  Query Match             33.9%;  Score 465;  DB 9;  Length 791;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-247;
  Matches  515;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy     351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      79 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 138

Qy     411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     139 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 198

Qy     471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     199 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 258

Qy     531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     259 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTTT 318

Qy     591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     319 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 378

Qy     651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     379 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 438

Qy     711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     439 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 498

Qy     771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     499 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 558

Qy     831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGA 866
           |||||||||||||||||||||||||||||||||||
Db     559 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGA 594
```

RESULT 10
BG339524
LOCUS BG339524 1006 bp mRNA linear EST 27-FEB-2001
DEFINITION 602437368F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4555072 5', mRNA sequence.
ACCESSION BG339524
VERSION BG339524.1 GI:13145962
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1006)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1255 row: h column: 17
High quality sequence stop: 539.
FEATURES Location/Qualifiers
source 1. .1006
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4555072"
/clone_lib="NIH_MGC_46"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 309 a 271 c 254 g 172 t
ORIGIN

Query Match 30.9%; Score 424; DB 10; Length 1006;
Best Local Similarity 100.0%; Pred. No. 8.9e-225;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     352 CAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCT 411
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118 CAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCT 177

Qy     412 AAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACC 471
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     178 AAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACC 237

Qy     472 ATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTT 531
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     238 ATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTT 297

Qy     532 GAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATG 591
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     298 GAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATG 357

Qy     592 GGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTG 651
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     358 GGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTG 417

Qy     652 GGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAA 711
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     418 GGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAA 477

Qy     712 AAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCAT 771
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     478 AAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCAT 537

Qy     772 GGAG 775
           ||||
Db     538 GGAG 541
```

RESULT 11
BE619147
LOCUS BE619147 796 bp mRNA linear EST 20-OCT-2000
DEFINITION 601472883F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3875733 5', mRNA sequence.
ACCESSION BE619147
VERSION BE619147.1 GI:9890085
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 796)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

```
            http://image.llnl.gov
            Plate: LLAM9635  row: f  column: 22
            High quality sequence stop: 671.
FEATURES              Location/Qualifiers
     source           1. .796
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:3875733"
                      /clone_lib="NIH_MGC_68"
                      /tissue_type="large cell carcinoma"
                      /lab_host="DH10B (phage-resistant)"
                      /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                      Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                      Average insert size 1.8 kb.   Library constructed by Life
                      Technologies. "
BASE COUNT      225 a    169 c    223 g    179 t
ORIGIN
```

```
  Query Match             30.3%;  Score 416;  DB 10;  Length 796;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-220;
  Matches  416;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     371 AGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGT 430
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     166 AGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGT 225

Qy     431 TTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCA 490
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     226 TTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCA 285

Qy     491 CCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCA 550
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     286 CCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCA 345

Qy     551 TCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAG 610
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     346 TCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAG 405

Qy     611 ACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCT 670
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     406 ACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCT 465

Qy     671 ATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACC 730
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     466 ATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACC 525

Qy     731 ATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTA 786
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     526 ATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTA 581
```

```
RESULT  12
BM194512
LOCUS       BM194512                497 bp    mRNA    linear   EST 13-DEC-2001
DEFINITION  TCAAP2Q11752 Pediatric pre-B cell acute lymphoblastic leukemia
            Baylor-HGSC project=TCBA Homo sapiens cDNA clone TCAAP1175, mRNA
            sequence.
ACCESSION   BM194512
VERSION     BM194512.1  GI:17653420
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 497)
  AUTHORS   Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr.,
            Gunaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
  TITLE     Pediatric Leukemia cDNA Sequencing Project (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Dr. Judith F. Margolin
            Texas Children's Cancer Center and Human Genome Sequencing Center
            at Baylor College of Medicine
            1102 Bates, MC3-3320 Houston, TX 77030, USA
            Tel: 832-824-4536
            Fax: 832-825-4038
            Email: clones@txccc.org
            Seq primer: M13 primer.
FEATURES              Location/Qualifiers
     source           1. .497
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="TCAAP1175"
                      /clone_lib="Pediatric pre-B cell acute lymphoblastic
                      leukemia Baylor-HGSC project=TCBA"
                      /sex="male"
                      /tissue_type="leukopheresis"
                      /cell_type="pre-B cell"
                      /dev_stage="pediatric 2 years"
                      /lab_host="DH10B"
                      /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
                      First strand cDNA was primed with an anchored
                      XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGGCCGCAGGAGGAG(T)VN
                      3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
                      was primed with a BamHI-dC primer
                      [5'AGAGAGCTCGGATCCGCGGCCGCAATAATAATAAT(C) 3'].
                      Double-stranded cDNA was then digested with BamHI and XhoI
                      and directionally cloned into the BamHI and SalI sites of
                      lambda pSB vector. Library went through one round of
                      normalization. Library was constructed by Wei Yu at RIKEN
                      of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
                      Itoh M, Nagaoka S, SasakiN, Okazaki Y, Muramatsu M,
                      Schneider C, Hayashizaki Y, High efficiency selection of
                      full-length cDNA by improved biotinylated cap trapper.,
                      DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT      143 a    118 c    129 g    107 t
ORIGIN
```

```
  Query Match             30.3%;  Score 415;  DB 10;  Length 497;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-220;
  Matches  415;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      80 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 139

Qy     411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     140 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 199

Qy     471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     200 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 259

Qy     531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     260 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 319

Qy     591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     320 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 379

Qy     651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     380 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 439

Qy     711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTA 765
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     440 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTA 494
```

```
RESULT  13
AU117167
LOCUS       AU117167                700 bp    mRNA    linear   EST 19-OCT-2000
DEFINITION  AU117167 HEMBA1 Homo sapiens cDNA clone HEMBA1000808 5', mRNA
```

```
            sequence.
ACCESSION   AU117167
VERSION     AU117167.1  GI:10932113
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 700)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .700
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEMBA1000808"
                     /clone_lib="HEMBA1"
                     /tissue_type="whole embryo, mainly head"
                     /dev_stage="embryo, 10 weeks"
                     /note="Vector: pME18SFL3"
BASE COUNT      188 a    158 c    154 g    197 t      3 others
ORIGIN
```

Query Match 30.1%; Score 413; DB 9; Length 700;
Best Local Similarity 99.7%; Pred. No. 1.2e-218;
Matches 583; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
Qy    787 ATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGAT 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGAT 60

Qy    847 AGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACC 906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACC 120

Qy    907 GAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCT 966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCT 180

Qy    967 CGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTAT 1026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTAT 240

Qy   1027 TACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTG 1086
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTG 300

Qy   1087 GAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGACAGAGT 1146
          |||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||||
Db    301 GAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCAGGGTGAATGACAGAGT 359

Qy   1147 TGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGATCCC 1206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 TGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGATCCC 419

Qy   1207 CTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCAGGGGTA 1266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 CTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCAGGGGTA 479
```

```
Qy   1267 CATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGCTCTTGA 1326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 CATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGCTCTTGA 539

Qy   1327 AATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
          |||||||||||||||||||||||||||||||||||||||||||||
Db    540 AATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 584
```

```
RESULT  14
BM147460
LOCUS       BM147460                485 bp    mRNA    linear   EST 30-NOV-2001
DEFINITION  TCAAP1Q11752 Pediatric acute myelogenous leukemia cell (FAB M1)
            Baylor-HGSC project=TCAA Homo sapiens cDNA clone TCAAP1175, mRNA
            sequence.
ACCESSION   BM147460
VERSION     BM147460.1  GI:17167317
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 485)
  AUTHORS   Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr.,
            Gunaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
  TITLE     Pediatric Leukemia cDNA Sequencing Project (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Dr. Judith F. Margolin
            Texas Children's Cancer Center and Human Genome Sequencing Center
            at Baylor College of Medicine
            1102 Bates, MC3-3320 Houston, TX 77030, USA
            Tel: 832-824-4536
            Fax: 832-825-4038
            Email: clones@txccc.org
            Seq primer: M13 primer.
FEATURES             Location/Qualifiers
     source          1. .485
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TCAAP1175"
                     /clone_lib="Pediatric acute myelogenous leukemia cell (FAB
                     M1) Baylor-HGSC project=TCAA"
                     /sex="male"
                     /tissue_type="leukopheresis"
                     /cell_type="myeloid cell"
                     /dev_stage="pediatric 6 years"
                     /lab_host="DH10B"
                     /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
                     First strand cDNA was primed with an anchored
                     XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGGCCGCAGGAGGAG(T)VN
                     3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
                     was primed with a BamHI-dC primer
                     [5'AGAGAGCTCGGATCCGCGGCCGCAATAATAATAAT(C) 3'].
                     Double-stranded cDNA was then digested with BamHI and XhoI
                     and directionally cloned into the BamHI and SalI sites of
                     lambda pSB vector. Library went through one round of
                     normalization. Library was constructed by Wei Yu at RIKEN
                     of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
                     Itoh M, Nagaoka S, SasakiN, Okazaki Y, Muramatsu M,
                     Schneider C, Hayashizaki Y, High efficiency selection of
                     full-length cDNA by improved biotinylated cap trapper.,
                     DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT      141 a    112 c    129 g    103 t
ORIGIN
```

Query Match 28.6%; Score 392; DB 10; Length 485;
Best Local Similarity 100.0%; Pred. No. 6e-207;
Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      80 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 139

Qy     411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     140 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 199

Qy     471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     200 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 259

Qy     531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     260 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 319

Qy     591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     320 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 379

Qy     651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     380 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 439

Qy     711 AAAAGCTTCTGAAGAGGACCATACAAATGCCG 742
           ||||||||||||||||||||||||||||||||
Db     440 AAAAGCTTCTGAAGAGGACCATACAAATGCCG 471
```

RESULT 15
BM150092

```
LOCUS       BM150092                532 bp    mRNA    linear   EST 30-NOV-2001
DEFINITION  TCAAP4Q11752 Pediatric acute myelogenous leukemia cell (FAB M1)
            Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAAP1175, mRNA
            sequence.
ACCESSION   BM150092
VERSION     BM150092.1  GI:17172294
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 532)
  AUTHORS   Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr.,
            Gunaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
  TITLE     Pediatric Leukemia cDNA Sequencing Project (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Dr. Judith F. Margolin
            Texas Children's Cancer Center and Human Genome Sequencing Center
            at Baylor College of Medicine
            1102 Bates, MC3-3320 Houston, TX 77030, USA
            Tel: 832-824-4536
            Fax: 832-825-4038
            Email: clones@txccc.org
            Seq primer: M13 primer.
FEATURES             Location/Qualifiers
     source          1. .532
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TCAAP1175"
                     /clone_lib="Pediatric acute myelogenous leukemia cell (FAB
                     M1) Baylor-HGSC project-TCAA"
                     /sex="male"
                     /tissue_type="leukopheresis"
                     /cell_type="myeloid cell"
                     /dev_stage="pediatric 6 years"
                     /lab_host="DH10B"
                     /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
                     First strand cDNA was primed with an anchored
                     XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGGCCGCAGGAGGAG(T)VN
                     3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
                     was primed with a BamH1-dC primer
                     [5'AGAGAGCTCGGATCCGCGGCCGCAATAATAATAAT(C) 3'].
                     Double-stranded cDNA was then digested with BamH1 and XhoI
                     and directionally cloned into the BamH1 and SalI sites of
                     lambda pSB vector. Library went through one round of
                     normalization. Library was constructed by Wei Yu at RIKEN
                     of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
                     Itoh M, Nagaoka S, SasakiN, Okazaki Y, Muramatsu M,
                     Schneider C, Hayashizaki Y, High efficiency selection of
                     full-length cDNA by improved biotinylated cap trapper.,
                     DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT      159 a    117 c    138 g    117 t      1 others
ORIGIN
```

```
  Query Match             28.6%;  Score 392;  DB 10;  Length 532;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-207;
  Matches  392;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     370 CAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCG 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      99 CAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCG 158

Qy     430 TTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     159 TTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 218

Qy     490 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     219 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGC 278

Qy     550 ATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACA 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     279 ATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACA 338

Qy     610 GACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTC 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     339 GACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTC 398

Qy     670 TATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGAC 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     399 TATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGAC 458

Qy     730 CATACAAATGCCGCCTGCTTCGCCTGCATCCT 761
           ||||||||||||||||||||||||||||||||
Db     459 CATACAAATGCCGCCTGCTTCGCCTGCATCCT 490
```

Search completed: August 16, 2002, 00:45:44
Job time: 6200 sec

```
OM of: US-09-613-508-2 to: Issued_Patents_NA:*   out_format : pfs

Date: Aug 15, 2002 10:16 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO_spool/US09613508/runat_14082002_112714_7793/app_query.fasta_1.701
 -DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=olig.rni
 -GAPOP=4.500 -GAPEXT=0.050 -MINMATCH=0.100 -LOOPCL=0.000
 -LOOPEXT=0.000 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=60.000
 -XGAPEXT=60.000 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000
 -YGAPEXT=60.000 -DELOP=6.000 -DELEXT=7.000 -START=1
 -MATRIX=oligo -TRANS=human40.cdi -LIST=45 -DOCALIGN=200
 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US09613508_@CGN1_1_54 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-613-508-2
Query length: 303
Database: Issued_Patents_NA:*
Database sequences: 383533
Database length: 122816752
Search time (sec): 64.210000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000


score_list:
Sequence          Strd Orig    ZScore   EScore Len  | Documentation  ..
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-1 +  189.00 3454.24  1.0e-184  1371  |
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-665-220-1 +   11.00  175.67  0.0426    1700  |
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-291-692-1 +   11.00  175.67  0.0426    1700  |
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-618-408B-1 +   11.00  170.37  0.0841    3535  |
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-852-936C-3 +   10.00  164.33  0.1825    639   |
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-300-328-3 +   10.00  164.33  0.1825    639   |  S
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-308-5 +   10.00  161.09  0.2767    1001  |
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-308-9 +   10.00  161.09  0.2767    1001  |
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-313-5 +   10.00  161.09  0.2767    1001  |
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-313-9 +   10.00  161.09  0.2767    1001  |
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-069-138-5 +   10.00  161.09  0.2767    1001  |
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-069-138-9 +   10.00  161.09  0.2767    1001  |
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-591-605-1 +   10.00  160.03  0.3167    1158  |
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-3 +   10.00  160.03  0.3169    1159  |
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-890-542A-1 +   10.00  159.92  0.3215    1177  |
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-890-542A-3 +   10.00  159.92  0.3215    1177  |
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-257-218-1 +   10.00  159.47  0.3402    1251  |
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-311-760-1 +   10.00  159.47  0.3402    1251  |
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-10 +   10.00  159.21  0.3518    1297  |
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-852-936C-2 +   10.00  157.80  0.4219    1578  |
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-300-328-2 +   10.00  157.80  0.4219    1578  |
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-3 +   10.00  154.09  0.6788    2635  |
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-773-608A-1 +   10.00  154.03  0.6843    2658  |
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-859-998-1322 -    9.00  168.54  0.1064    28  |
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-225-928-1322 -    9.00  168.54  0.1064    28  |
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-368-704C-34 +    9.00  136.11  6.81      2482  |
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-258-287B-34 +    9.00  136.11  6.81      2483  |
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-394-189B-1 +    9.00  127.97  19.35     7653  |
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-258-287B-1 +    9.00  127.97  19.35     7653  |
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-368-704C-1 +    9.00  127.97  19.35     7653  |
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US93-05701-18 +    9.00  127.97  19.35     7653
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US93-05705-1 +    9.00  127.97  19.35     7653
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-859-998-1321 +    8.00  150.66  1.05      26  |
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-225-928-1321 +    8.00  150.66  1.05      26  |
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-258-287B-45 +    8.00  139.60  4.35      120  |
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-258-287B-47 +    8.00  139.60  4.35      120  |
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-258-287B-49 +    8.00  139.60  4.35      120  |
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-368-704C-44 +    8.00  139.60  4.35      120  |
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-368-704C-46 +    8.00  139.60  4.35      120
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-368-704C-48 +    8.00  139.60  4.35      120
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-247-155-53 +    8.00  130.13  14.67     445
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US94-07127A-3 +    8.00  121.69  43.30     1
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-368-704C-50 +    8.00  121.56  44.03     1456
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-368-704C-52 +    8.00  118.68  63.67     2167
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-258-287B-54 +    8.00  118.66  63.86     2174

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-1

seq_documentation_block:
; Sequence 1, Application US/08462969B
; Patent No. 6087150
;  GENERAL INFORMATION:
;    APPLICANT:  He, Wei-Wu et al.
;    TITLE OF INVENTION:  Interleukin-1 Beta Converting Enzyme
;    TITLE OF INVENTION:  Like Apoptosis Protease 3 and 4
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Ave.
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,969B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/334,251
;      FILING DATE:  11-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF140P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1371 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-462-969B-1


alignment_scores:
            Quality:  189.00              Length:    189
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x US-08-462-969B-1   ..

 Align seg 1/1  to: US-08-462-969B-1  from: 1  to: 1371

       5 GlnGlyCysIleGluGluGlnGlyValGluAspSerAlaAsnGluAspSe 21
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     352 CAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTC 401

      21 rValAspAlaLysProAspArgSerSerPheValProSerLeuPheSerL 38
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     402 AGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTA 451

      38 ysLysLysLysAsnValThrMetArgSerIleLysThrThrArgAspArg 54
```

```
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 452 AGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGA 501

  55 ValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGlyLysCysIl 71
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 502 GTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCAT 551

  71 eIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGlyValArgA 88
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 552 CATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAA 601

  88 snGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheArgSerLeu 104
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 602 ACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTG 651

 105 GlyPheAspValIleValTyrAsnAspCysSerCysAlaLysMetGlnAs 121
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 652 GGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGA 701

 121 pLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAlaCysPheA 138
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 702 TCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCG 751

 138 laCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrGlyLysAsp 154
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 752 CCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGAT 801

 155 GlyValThrProIleLysAspLeuThrAlaHisPheArgGlyAspArgCy 171
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 802 GGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATG 851

 171 sLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAlaCysArgG 188
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 852 CAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAG 901

 188 lyThrGluLeuAspAsp 193
     |||||||||||||||||
 902 GGACCGAGCTTGATGAT 918
```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-665-220-1

```
seq_documentation_block:
; Sequence 1, Application US/08665220
; Patent No. 5786173
;  GENERAL INFORMATION:
;    APPLICANT:  Alnemri, Emad S.
;    APPLICANT:  Fernandes-Alnemri, Teresa
;    APPLICANT:  Litwack, Gerald
;    APPLICANT:  Armstrong, Robert
;    APPLICANT:  Tomaselli, Kevin
;    TITLE OF INVENTION:  Mch4 and Mch5, Apoptotic Proteases,
;    TITLE OF INVENTION:  Nucleic Acids Encoding and Methods of Use
;    NUMBER OF SEQUENCES:  70
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/665,220
;      FILING DATE:  14-JUN-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/618,408
;      FILING DATE:  19-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-ID 2165
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1700 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  148..1584
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..3535
;      OTHER INFORMATION:  /note= "Mch4"
US-08-665-220-1


alignment_scores:
            Quality:   11.00              Length:     11
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x US-08-665-220-1    ..

 Align seg 1/1  to: US-08-665-220-1  from: 1  to: 1700

     176 GluLysProLysLeuPhePheIleGlnAlaCys 186
         |||||||||||||||||||||||||||||||
    1189 GAAAAACCTAAACTCTTTTTCATCCAGGCCTGC 1221

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-291-692-1

seq_documentation_block:
; Sequence 1, Application US/09291692
; Patent No. 6287795
;  GENERAL INFORMATION:
;    APPLICANT:  Alnemri, Emad S.
;    APPLICANT:  Fernandes-Alnemri, Teresa
;    APPLICANT:  Litwack, Gerald
;    APPLICANT:  Armstrong, Robert
;    APPLICANT:  Tomaselli, Kevin
;    TITLE OF INVENTION:  MCH4 AND MCH5, APOPTOTIC PROTEASE,
;    TITLE OF INVENTION:  NUCLEIC ACIDS ENCODING AND METHODS OF USE
;    NUMBER OF SEQUENCES:  75
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USe
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/291,692
;      FILING DATE:  04-13-1999
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Christiansen, william T.
;      REGISTRATION NUMBER:  44,614
;      REFERENCE/DOCKET NUMBER:  480140.424C3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
```

```
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1700 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  148..1584
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..1700
;      OTHER INFORMATION:  /note= "Mch4"
US-09-291-692-1


alignment_scores:
            Quality:   11.00              Length:     11
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x US-09-291-692-1   ..

 Align seg 1/1  to: US-09-291-692-1  from: 1  to: 1700

     176 GluLysProLysLeuPhePheIleGlnAlaCys 186
         ||||||||||||||||||||||||||||||||||
    1189 GAAAAACCTAAACTCTTTTTCATCCAGGCCTGC 1221

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-618-408B-1

seq_documentation_block:
; Sequence 1, Application US/08618408B
; Patent No. 5851815
;  GENERAL INFORMATION:
;    APPLICANT:  Alnemri, Emad S.
;    APPLICANT:  Fernandes-Alnemri, Teresa
;    APPLICANT:  Litwack, Gerald
;    APPLICANT:  Armstrong, Robert
;    APPLICANT:  Tomaselli, Kevin
;    TITLE OF INVENTION:  Mch4 and Mch5, No. 5851815el Apoptotic
;    TITLE OF INVENTION:  Proteases, Nucleic Acids Encoding and Methods of Use
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/618,408B
;      FILING DATE:  19-MAR-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-ID 1957
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3535 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  546..1584
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..3535
;      OTHER INFORMATION:  /note= "Mch4"
US-08-618-408B-1


alignment_scores:
            Quality:   11.00              Length:     11
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x US-08-618-408B-1   ..

 Align seg 1/1  to: US-08-618-408B-1  from: 1  to: 3535

     176 GluLysProLysLeuPhePheIleGlnAlaCys 186
         ||||||||||||||||||||||||||||||||||
    1188 GAAAAACCTAAACTCTTTTTCATCCAGGCCTGC 1220

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-852-936C-3

seq_documentation_block:
; Sequence 3, Application US/08852936C
; Patent No. 6010878
;  GENERAL INFORMATION:
;    APPLICANT:  DIXIT, VISHVA M.
;    APPLICANT:  HE, WEI-WU
;    APPLICANT:  KIKLY, KRISTINE K.
;    APPLICANT:  RUBEN, STEVEN M.
;    TITLE OF INVENTION:  INTERLEUKIN-1 BETA CONVERTING
;    TITLE OF INVENTION:  ENZYME LIKE APOPTOTIC PROTEASE-6
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ratner & Prestia
;      STREET:  P.O. Box 980
;      CITY:  Valley Forge
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/852,936C
;      FILING DATE:  08-MAY-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/018,961
;      FILING DATE:  05-JUN-1996
;      APPLICATION NUMBER:  60/020,344
;      FILING DATE:  23-MAY-1996
;      APPLICATION NUMBER:  60/017,949
;      FILING DATE:  20-MAY-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Prestia, Paul F
;      REGISTRATION NUMBER:  23,031
;      REFERENCE/DOCKET NUMBER:  p50483-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-407-0700
;      TELEFAX:  610-407-0700
;      TELEX:  846169
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  639 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-852-936C-3


alignment_scores:
            Quality:   10.00              Length:     10
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x US-08-852-936C-3   ..

 Align seg 1/1  to: US-08-852-936C-3  from: 1  to: 639

     177 LysProLysLeuPhePheIleGlnAlaCys 186
         ||||||||||||||||||||||||||||||
     214 AAGCCCAAGCTCTTTTTCATCCAGGCCTGT 243

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-300-328-3

seq_documentation_block:
; Sequence 3, Application US/09300328
; Patent No. 6294169
;  GENERAL INFORMATION:
;    APPLICANT:  DIXIT, VISHVA M.
;    APPLICANT:  HE, WEI-WU
;    APPLICANT:  KIKLY, KRISTINE K.
;    APPLICANT:  RUBEN, STEVEN M.
;    TITLE OF INVENTION:  INTERLEUKIN-1 BETA CONVERTING
;    TITLE OF INVENTION:  ENZYME LIKE APOPTOTIC PROTEASE-6
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ratner & Prestia
;      STREET:  P.O. Box 980
;      CITY:  Valley Forge
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/300,328
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/852,936
;      FILING DATE:  08-MAY-1997
;      APPLICATION NUMBER:  60/018,961
;      FILING DATE:  05-JUN-1996
;      APPLICATION NUMBER:  60/020,344
;      FILING DATE:  23-MAY-1996
;      APPLICATION NUMBER:  60/017,949
;      FILING DATE:  20-MAY-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Prestia, Paul F
;      REGISTRATION NUMBER:  23,031
;      REFERENCE/DOCKET NUMBER:  p50483-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-407-0700
;      TELEFAX:  610-407-0700
;      TELEX:  846169
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  639 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-300-328-3


alignment_scores:
            Quality:   10.00              Length:     10
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x US-09-300-328-3   ..

 Align seg 1/1  to: US-09-300-328-3  from: 1  to: 639

     177 LysProLysLeuPhePheIleGlnAlaCys 186
         ||||||||||||||||||||||||||||||
     214 AAGCCCAAGCTCTTTTTCATCCAGGCCTGT 243

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-308-5

seq_documentation_block:
; Sequence 5, Application US/08964308
; Patent No. 6066715
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBERT J. NORTH - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,308
;      FILING DATE:  04-NOV-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  NORTH, ROBERT J
;      REGISTRATION NUMBER:  27,366
;      REFERENCE/DOCKET NUMBER:  19840 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-7262
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-308-5


alignment_scores:
            Quality:   10.00              Length:     10
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x US-08-964-308-5   ..

 Align seg 1/1  to: US-08-964-308-5  from: 1  to: 1001

     183 IleGlnAlaCysArgGlyThrGluLeuAsp 192
```

```
                 ||||||||||||||||||||||||||||||
     541 ATTCAGGCCTGCCGTGGTACAGAACTGGAC 570

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-308-9

seq_documentation_block:
; Sequence 9, Application US/08964308
; Patent No. 6066715
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBERT J. NORTH - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,308
;      FILING DATE:  04-NOV-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  NORTH, ROBERT J
;      REGISTRATION NUMBER:  27,366
;      REFERENCE/DOCKET NUMBER:  19840 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-7262
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-308-9


alignment_scores:
            Quality:   10.00              Length:     10
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x US-08-964-308-9   ..

 Align seg 1/1  to: US-08-964-308-9  from: 1  to: 1001

     292 ValSerMetLeuThrLysGluLeuTyrPhe 301
         ||||||||||||||||||||||||||||||
     859 GTTTCCATGCTCACAAAAGAACTCTATTTT 888

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-313-5

seq_documentation_block:
; Sequence 5, Application US/08964313
; Patent No. 6114132
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  GRESSER, MICHAEL
;    APPLICANT:  KENNEDY, BRIAN
;    APPLICANT:  NICHOLSON, DONALD
;    APPLICANT:  RAMACHANDRAN, CHIDAMBARAN
;    APPLICANT:  SKOREY, KATHRYN
;    APPLICANT:  FORD-HUTCHINSON, ANTHONY
;    TITLE OF INVENTION:  PHOSPHATASE BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,313
;      FILING DATE:  04-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/030,408
;      FILING DATE:  04-NOV-1996
;      APPLICATION NUMBER:  PCT/CA97/00825
;      FILING DATE:  03 -NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19824Y
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-313-5


alignment_scores:
            Quality:   10.00              Length:     10
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x US-08-964-313-5   ..

 Align seg 1/1  to: US-08-964-313-5  from: 1  to: 1001

     183 IleGlnAlaCysArgGlyThrGluLeuAsp 192
         ||||||||||||||||||||||||||||||
     541 ATTCAGGCCTGCCGTGGTACAGAACTGGAC 570

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-313-9

seq_documentation_block:
; Sequence 9, Application US/08964313
; Patent No. 6114132
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  GRESSER, MICHAEL
;    APPLICANT:  KENNEDY, BRIAN
;    APPLICANT:  NICHOLSON, DONALD
;    APPLICANT:  RAMACHANDRAN, CHIDAMBARAN
;    APPLICANT:  SKOREY, KATHRYN
;    APPLICANT:  FORD-HUTCHINSON, ANTHONY
```

```
;    TITLE OF INVENTION:  PHOSPHATASE BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:-
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,313
;      FILING DATE:  04-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/030,408
;      FILING DATE:  04-NOV-1996
;      APPLICATION NUMBER:  PCT/CA97/00825
;      FILING DATE:  03 -NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19824Y
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-313-9


alignment_scores:
            Quality:   10.00              Length:     10
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x US-08-964-313-9    ..

 Align seg 1/1  to: US-08-964-313-9  from: 1  to: 1001

     292 ValSerMetLeuThrLysGluLeuTyrPhe 301
         |||||||||||||||||||||||||||||
     859 GTTTCCATGCTCACAAAAGAACTCTATTTT 888

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-069-138-5

seq_documentation_block:
; Sequence 5, Application US/09069138
; Patent No. 6348572
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  DUFRESNE, CLAUDE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  LEBLANC, YVES
;    APPLICANT:  ROY, PATRICK
;    APPLICANT:  YOUNG, ROBERT N.
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/069,138
;      FILING DATE:  29-APR-1998
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19840YIA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-069-138-5


alignment_scores:
            Quality:   10.00              Length:     10
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x US-09-069-138-5    ..

 Align seg 1/1  to: US-09-069-138-5  from: 1  to: 1001

     183 IleGlnAlaCysArgGlyThrGluLeuAsp 192
         |||||||||||||||||||||||||||||||
     541 ATTCAGGCCTGCCGTGGTACAGAACTGGAC 570

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-069-138-9

seq_documentation_block:
; Sequence 9, Application US/09069138
; Patent No. 6348572
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  DUFRESNE, CLAUDE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  LEBLANC, YVES
;    APPLICANT:  ROY, PATRICK
;    APPLICANT:  YOUNG, ROBERT N.
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/069,138
;      FILING DATE:  29-APR-1998
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19840YIA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-069-138-9


alignment_scores:
            Quality:   10.00              Length:     10
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x US-09-069-138-9    ..

 Align seg 1/1  to: US-09-069-138-9  from: 1  to: 1001

     292 ValSerMetLeuThrLysGluLeuTyrPhe 301
         |||||||||||||||||||||||||||||
     859 GTTTCCATGCTCACAAAAGAACTCTATTTT 888

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-591-605-1

seq_documentation_block:
; Sequence 1, Application US/08591605
; Patent No. 6060238
;  GENERAL INFORMATION:
;    APPLICANT:  Dixit, Vishva M.
;    TITLE OF INVENTION:  METHOD AND COMPOSITION FOR REGULATING
;    TITLE OF INVENTION:  APOPTOSIS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 PAGE MILL ROAD
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/591,605
;      FILING DATE:  09-FEB-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KONSKI, ANTOINETTE F.
;      REGISTRATION NUMBER:  34,202
;      REFERENCE/DOCKET NUMBER:  20344-21036.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  706141 MRSNFOERS SFO
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1158 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  224..1054
;      OTHER INFORMATION:  /product= "Yama peptide sequence"
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  704..718
;      OTHER INFORMATION:  /note= "region coding for
;      OTHER INFORMATION:  pentapeptide, wherein cysteine is catalytic"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "atg")
;      OTHER INFORMATION:  /note= "Mutation of catalytic cysteine to methionine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gcc")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gcg")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gct")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gca")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
US-08-591-605-1


alignment_scores:
            Quality:   10.00              Length:     10
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x US-08-591-605-1    ..

 Align seg 1/1  to: US-08-591-605-1  from: 1  to: 1158

     183 IleGlnAlaCysArgGlyThrGluLeuAsp 192
         ||||||||||||||||||||||||||||||
     701 ATTCAGGCCTGCCGTGGTACAGAACTGGAC 730

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-3

seq_documentation_block:
; Sequence 3, Application US/08462969B
; Patent No. 6087150
;  GENERAL INFORMATION:
;    APPLICANT:  He, Wei-Wu et al.
;    TITLE OF INVENTION:  Interleukin-1 Beta Converting Enzyme
;    TITLE OF INVENTION:  Like Apoptosis Protease 3 and 4
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Ave.
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US/08/462,969B
;       FILING DATE:  05-JUN-1995
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/334,251
;       FILING DATE:  11-NOV-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Brookes, A. Anders
;       REGISTRATION NUMBER:  36,373
;       REFERENCE/DOCKET NUMBER:  PF140P1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  301-309-8504
;       TELEFAX:  301-309-8439
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1159 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-462-969B-3


alignment_scores:
            Quality:   10.00              Length:     10
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x US-08-462-969B-3    ..

 Align seg 1/1  to: US-08-462-969B-3  from: 1  to: 1159

     183 IleGlnAlaCysArgGlyThrGluLeuAsp 192
         ||||||||||||||||||||||||||||||
     700 ATTCAGGCCTGCCGTGGTACAGAACTGGAC 729

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-890-542A-1

seq_documentation_block:
; Sequence 1, Application US/08890542A
; Patent No. 5840509
;  GENERAL INFORMATION:
;     APPLICANT:  Ni, Binhui
;     APPLICANT:  Paul, Steven M
;     APPLICANT:  Wu, Xin
;     TITLE OF INVENTION:  PROTEASE AND RELATED NUCLEIC ACID
;     TITLE OF INVENTION:  COMPOUNDS
;     NUMBER OF SEQUENCES:  3
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Eli Lilly and Company
;       STREET:  Lilly Corporate Center
;       CITY:  Indianapolis
;       STATE:  Indiana
;       COUNTRY:  United States of America
;       ZIP:  46285
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/890,542A
;       FILING DATE:  09-JUL-1997
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Gaylo, Paul J.
;       REGISTRATION NUMBER:  36,808
;       REFERENCE/DOCKET NUMBER:  X-10704
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (317) 276-0756
;       TELEFAX:  (317) 276-3861
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1177 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  76..906
US-08-890-542A-1


alignment_scores:
            Quality:   10.00              Length:     10
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x US-08-890-542A-1    ..

 Align seg 1/1  to: US-08-890-542A-1  from: 1  to: 1177

     183 IleGlnAlaCysArgGlyThrGluLeuAsp 192
         ||||||||||||||||||||||||||||||
     553 ATTCAGGCCTGCCGAGGTACAGAGCTGGAC 582
```

```
OM of: US-09-613-508-4 to: Issued_Patents_NA:*   out_format : pfs

Date: Aug 15, 2002 10:16 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=x1h
 -Q=/cgn2_1/USPTO_spool/US09613508/runat_14082002_112714_7793/app_query.fasta_1.701
 -DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=olig.rni
 -GAPOP=4.500 -GAPEXT=0.050 -MINMATCH=0.100 -LOOPCL=0.000
 -LOOPEXT=0.000 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=60.000
 -XGAPEXT=60.000 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000
 -YGAPEXT=60.000 -DELOP=6.000 -DELEXT=7.000 -START=1
 -MATRIX=oligo -TRANS=human40.cdi -LIST=45 -DOCALIGN=200
 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US09613508_@CGN1_1_54 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-613-508-4
Query length: 277
Database: Issued_Patents_NA:*
Database sequences: 383533
Database length: 122816752
Search time (sec): 64.210000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000


score_list:
Sequence                Strd Orig    %Score   EScore Len  ! Documentation  ..
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-308-5 +  277.00 5158.11  1.3e-279  1001  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-313-5 +  277.00 5158.11  1.3e-279  1001  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-069-138-5 +  277.00 5158.11  1.3e-279  1001  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-591-605-1 +  277.00 5157.06  1.5e-279  1158  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-3 +  277.00 5157.05  1.5e-279  1159  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-3 +  189.00 3505.09  1.5e-187  2635  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-308-9 +  176.00 3268.91  2.2e-174  1001  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-313-9 +  176.00 3268.91  2.2e-174  1001  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-069-138-9 +  176.00 3268.91  2.2e-174  1001  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-10 +   35.00  629.63  2.2e-27   1297  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-890-542A-1 +   30.00  536.81  3.3e-22   1177  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-890-542A-3 +   30.00  536.81  3.3e-22   1177  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-800-007A-1 +   11.00  203.77  0.0012    53  !  S
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-800-007A-5 +   11.00  203.77  0.0012    53  !  S
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-208-966-24 +   11.00  200.36  0.0018    85  !  S
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-208-966-36 +   11.00  200.36  0.0018    85  !  S
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-208-966-20 +   10.00  181.25  0.0208    90  !  S
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-208-966-32 +   10.00  181.25  0.0208    90  !  S
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-1 +   10.00  161.61  0.2588  1371  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-800-007A-3 +    9.00  167.23  0.1259    47  !  S
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-800-007A-7 +    9.00  167.23  0.1259    47  !  S
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-800-007A-4 -    9.00  166.64  0.1357    51  !  S
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-800-007A-8 -    9.00  166.64  0.1357    51  !  S
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-800-007A-2 -    9.00  166.50  0.1382    52  !  S
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-800-007A-6 -    9.00  166.50  0.1382    52  !  S
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-11 +    8.00  153.07  0.7730    25  !  S
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-6 +    8.00  152.52  0.8300    27  !  Se
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-522-813-1 +    8.00  125.26  27.38     1183  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-446-925-6 +    8.00  124.51  30.15     1313  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-146-331-6 +    8.00  124.51  30.15     1313  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-896-885-6 +    8.00  124.51  30.15     1313  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-375-256-6 +    8.00  124.51  30.15     1313  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-446-925-4 +    8.00  123.33  35.05     1545  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-146-331-4 +    8.00  123.33  35.05     1545  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-896-885-4 +    8.00  123.33  35.05     1545  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-375-256-4 +    8.00  123.33  35.05     1545  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-261-822A-9 +    8.00  122.55  38.75     1722  !
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-07744A-9 +    8.00  122.55  38.75     1722
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-248-021A-1 -    8.00  120.06  53.35     2433
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-4 +    7.00  135.63  7.24      21  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-13 +    7.00  134.66  8.20      24  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-7 +    7.00  132.82  10.39     31  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-12 -    7.00  128.95  17.06     53
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-208-966-21 +    7.00  126.74  22.64     72  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-208-966-34 +    7.00  126.74  22.64     72  !

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-308-5

seq_documentation_block:
; Sequence 5, Application US/08964308
; Patent No. 6066715
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBERT J. NORTH - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,308
;      FILING DATE:  04-NOV-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  NORTH, ROBERT J
;      REGISTRATION NUMBER:  27,366
;      REFERENCE/DOCKET NUMBER:  19840 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-7262
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-308-5


alignment_scores:
            Quality:  277.00              Length:    277
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x US-08-964-308-5   ..

 Align seg 1/1  to: US-08-964-308-5  from: 1  to: 1001

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||
164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

 51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
     ||||||||||||||||||||||||||||||||||||||||||||||||||
214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

 67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
     ||||||||||||||||||||||||||||||||||||||||||||||||||
264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

 84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
     ||||||||||||||||||||||||||||||||||||||||||||||||||
364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
     ||||||||||||||||||||||||||||||||||||||||||||||||||
414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 563

167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
     ||||||||||||||||||||||||||||||||||||||||||||||||||
564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
     ||||||||||||||||||||||||||||||||||||||||||||||||||
714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
     ||||||||||||||||||||||||||||||
864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894
```

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-313-5

```
seq_documentation_block:
; Sequence 5, Application US/08964313
; Patent No. 6114132
;  GENERAL INFORMATION:
;     APPLICANT:  DESMARAIS, SYLVIE
;     APPLICANT:  FRIESEN, RICHARD
;     APPLICANT:  GRESSER, MICHAEL
;     APPLICANT:  KENNEDY, BRIAN
;     APPLICANT:  NICHOLSON, DONALD
;     APPLICANT:  RAMACHANDRAN, CHIDAMBARAN
;     APPLICANT:  SKOREY, KATHRYN
;     APPLICANT:  FORD-HUTCHINSON, ANTHONY
;     TITLE OF INVENTION:  PHOSPHATASE BINDING ASSAY
;     NUMBER OF SEQUENCES:  15
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;       STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;       CITY:  RAHWAY
;       STATE:  NJ
;       COUNTRY:  USA
;       ZIP:  07065
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/964,313
;       FILING DATE:  04-NOV-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  60/030,408
;       FILING DATE:  04-NOV-1996
;       APPLICATION NUMBER:  PCT/CA97/00825
;       FILING DATE:  03 -NOV-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  DURETTE, PHILIPPE L.
;       REGISTRATION NUMBER:  35,125
;       REFERENCE/DOCKET NUMBER:  19824Y
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  732-594-4568
;       TELEFAX:  732-594-4720
;       TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1001 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-964-313-5


alignment_scores:
            Quality:  277.00              Length:    277
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x US-08-964-313-5   ..

 Align seg 1/1  to: US-08-964-313-5  from: 1  to: 1001

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
      314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

      101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

      117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

      134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

      151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 563

      167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

      184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

      201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

      217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

      234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

      251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

      267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
          ||||||||||||||||||||||||||||||
      864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-069-138-5

seq_documentation_block:
; Sequence 5, Application US/09069138
; Patent No. 6348572
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  DUFRESNE, CLAUDE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  LEBLANC, YVES
;    APPLICANT:  ROY, PATRICK
;    APPLICANT:  YOUNG, ROBERT N.
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/069,138
;      FILING DATE:  29-APR-1998
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19840YIA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-069-138-5


alignment_scores:
            Quality:  277.00              Length:    277
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x US-09-069-138-5   ..

 Align seg 1/1  to: US-09-069-138-5  from: 1  to: 1001

        1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
          ||||||||||||||||||||||||||||||||||||||||||||||||||
       64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

       17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

       34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

       51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

       67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

       84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

      101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

      117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

      134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

      151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 563
```

```
     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-591-605-1

seq_documentation_block:
; Sequence 1, Application US/08591605
; Patent No. 6060238
;  GENERAL INFORMATION:
;    APPLICANT:  Dixit, Vishva M.
;    TITLE OF INVENTION:  METHOD AND COMPOSITION FOR REGULATING
;    TITLE OF INVENTION:  APOPTOSIS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 PAGE MILL ROAD
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/591,605
;      FILING DATE:  09-FEB-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KONSKI, ANTOINETTE F.
;      REGISTRATION NUMBER:  34,202
;      REFERENCE/DOCKET NUMBER:  20344-21036.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  706141 MRSNFOERS SFO
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1158 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  224..1054
;      OTHER INFORMATION:  /product= "Yama peptide sequence"
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  704..718
;      OTHER INFORMATION:  /note= "region coding for
;      OTHER INFORMATION:  pentapeptide, wherein cysteine is catalytic"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "atg")
;      OTHER INFORMATION:  /note= "Mutation of catalytic cysteine to methionine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gcc")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gcg")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gct")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gca")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
US-08-591-605-1


alignment_scores:
            Quality:  277.00              Length:    277
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x US-08-591-605-1    ..

 Align seg 1/1  to: US-08-591-605-1  from: 1  to: 1158

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     224 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTGGA 273

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     274 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 323

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     324 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 373

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     374 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 423

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     424 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 473

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     474 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 523

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     524 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 573

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     574 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 623

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     624 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 673
```

```
     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     674 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 723

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     724 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 773

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     774 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 823

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     824 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 873

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     874 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 923

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     924 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 973

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     974 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1023

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
    1024 CATGCTCACAAAAGAACTCTATTTTTATCAC 1054

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-3

seq_documentation_block:
; Sequence 3, Application US/08462969B
; Patent No. 6087150
;  GENERAL INFORMATION:
;    APPLICANT:  He, Wei-Wu et al.
;    TITLE OF INVENTION:  Interleukin-1 Beta Converting Enzyme
;    TITLE OF INVENTION:  Like Apoptosis Protease 3 and 4
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Ave.
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,969B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/334,251
;      FILING DATE:  11-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF140P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1159 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-462-969B-3


alignment_scores:
            Quality:  277.00              Length:    277
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x US-08-462-969B-3   ..

 Align seg 1/1  to: US-08-462-969B-3  from: 1  to: 1159

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     223 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 272

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     273 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 322

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     323 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 372

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     373 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 422

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     423 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 472

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     473 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 522

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     523 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 572

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     573 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 622

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     623 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 672

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     673 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 722

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     723 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 772

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     773 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 822

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     823 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 872

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     873 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 922

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     923 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 972

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     973 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1022

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
    1023 CATGCTCACAAAAGAACTCTATTTTTATCAC 1053

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-3

seq_documentation_block:
; Sequence 3, Application US/09484617
; Patent No. 6303374
; GENERAL INFORMATION:
;  APPLICANT: Hong Zhang
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF CASPASE 3 EXPRESSION
;  FILE REFERENCE: RTS-0103
;  CURRENT APPLICATION NUMBER: US/09/484,617
;  CURRENT FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 176
; SEQ ID NO 3
;   LENGTH: 2635
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (225)...(1058)
US-09-484-617-3


alignment_scores:
            Quality:  189.00              Length:    189
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x US-09-484-617-3   ..

 Align seg 1/1  to: US-09-484-617-3  from: 1  to: 2635

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     225 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 274

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     275 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 324

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     325 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 424

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     425 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 474

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     475 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 524

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     525 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 574

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     575 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 624

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     625 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 674

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 724

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     725 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 774

     184 ysHisLysIleProVal 189
         |||||||||||||||||
     775 GTCATAAAATACCAGTG 791

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-308-9

seq_documentation_block:
; Sequence 9, Application US/08964308
; Patent No. 6066715
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBERT J. NORTH - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,308
;      FILING DATE:  04-NOV-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  NORTH, ROBERT J
;      REGISTRATION NUMBER:  27,366
;      REFERENCE/DOCKET NUMBER:  19840 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-7262
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-308-9


alignment_scores:
            Quality:  176.00              Length:    277
              Ratio:  0.638                 Gaps:      0
 Percent Similarity:  99.639   Percent Identity: 99.639

alignment_block:
 US-09-613-508-4 x US-08-964-308-9   ..

 Align seg 1/1  to: US-08-964-308-9  from: 1  to: 1001
```

```
      1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

     17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

     34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

     51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        ||||||||||||||||||||||||||||||||||  ||||||||||||||
    514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTCCCGTGGTACAGA 563

    167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

    184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

    201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

    217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

    234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

    251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

    267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
        |||||||||||||||||||||||||||||||
    864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-313-9

seq_documentation_block:
; Sequence 9, Application US/08964313
; Patent No. 6114132
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  GRESSER, MICHAEL
;    APPLICANT:  KENNEDY, BRIAN
;    APPLICANT:  NICHOLSON, DONALD
;    APPLICANT:  RAMACHANDRAN, CHIDAMBARAN
;    APPLICANT:  SKOREY, KATHRYN
;    APPLICANT:  FORD-HUTCHINSON, ANTHONY
;    TITLE OF INVENTION:  PHOSPHATASE BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,313
;      FILING DATE:  04-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/030,408
;      FILING DATE:  04-NOV-1996
;      APPLICATION NUMBER:  PCT/CA97/00825
;      FILING DATE:  03 -NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19824Y
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-313-9


alignment_scores:
           Quality:  176.00              Length:    277
             Ratio:  0.638                 Gaps:      0
 Percent Similarity:  99.639   Percent Identity: 99.639

alignment_block:
 US-09-613-508-4 x US-08-964-313-9   ..

 Align seg 1/1  to: US-08-964-313-9  from: 1  to: 1001

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213
```

```
     51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        ||||||||||||||||||||||||||||||||||  ||||||||||||||
    514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTCCCGTGGTACAGA 563

    167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

    184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

    201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

    217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

    234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

    251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

    267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
        ||||||||||||||||||||||||||||||
    864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-069-138-9

seq_documentation_block:
; Sequence 9, Application US/09069138
; Patent No. 6348572
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  DUFRESNE, CLAUDE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  LEBLANC, YVES
;    APPLICANT:  ROY, PATRICK
;    APPLICANT:  YOUNG, ROBERT N.
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/069,138
;      FILING DATE:  29-APR-1998
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19840YIA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-069-138-9


alignment_scores:
            Quality:  176.00              Length:    277
              Ratio:  0.638                 Gaps:      0
 Percent Similarity:  99.639   Percent Identity: 99.639

alignment_block:
 US-09-613-508-4 x US-09-069-138-9   ..

 Align seg 1/1  to: US-09-069-138-9  from: 1  to: 1001

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     464 TTGACCTGAAAAAAATAACAAACTTTTCAGAGGGGATCGTTGTAGAAGT 513

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         |||||||||||||||||||||||||||||||||   ||||||||||
     514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTCCCGTGGTACAGA 563

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         |||||||||||||||||||||||||||||||||||||||||||||||||
     664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         |||||||||||||||||||||||||||||
     864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894
```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-10

```
seq_documentation_block:
; Sequence 10, Application US/09484617
; Patent No. 6303374
; GENERAL INFORMATION:
;  APPLICANT: Hong Zhang
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF CASPASE 3 EXPRESSION
;  FILE REFERENCE: RTS-0103
;  CURRENT APPLICATION NUMBER: US/09/484,617
;  CURRENT FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 176
; SEQ ID NO 10
;   LENGTH: 1297
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (68)...(901)
US-09-484-617-10


alignment_scores:
            Quality:   35.00              Length:     35
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x US-09-484-617-10   ..

 Align seg 1/1  to: US-09-484-617-10  from: 1  to: 1297

     186 LysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGl 202
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     623 AAGATACCGGTGGAGGCTGACTTCCTGTATGCTTACTCTACAGCACCTGG 672

     202 yTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerL 219
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     673 TTACTATTCCTGGAGAAATTCAAAGGACGGGTCGTGGTTCATCCAGTCCC 722

     219 euCys 220
         |||||
     723 TTTGC 727
```

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-890-542A-1

```
seq_documentation_block:
; Sequence 1, Application US/08890542A
; Patent No. 5840509
;  GENERAL INFORMATION:
;    APPLICANT:  Ni, Binhui
;    APPLICANT:  Paul, Steven M
;    APPLICANT:  Wu, Xin
;    TITLE OF INVENTION:  PROTEASE AND RELATED NUCLEIC ACID
;    TITLE OF INVENTION:  COMPOUNDS
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  United States of America
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/890,542A
;      FILING DATE:  09-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gaylo, Paul J.
;      REGISTRATION NUMBER:  36,808
;      REFERENCE/DOCKET NUMBER:  X-10704
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (317) 276-0756
;      TELEFAX:  (317) 276-3861
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1177 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  76..906
US-08-890-542A-1


alignment_scores:
            Quality:   30.00              Length:     30
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x US-08-890-542A-1   ..

 Align seg 1/1  to: US-08-890-542A-1  from: 1  to: 1177

     148 CysArgSerLeuThrGlyLysProLysLeuPheIleIleGlnAlaCysAr 164
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     517 TGCCGGAGTCTGACTGGAAAGCCGAAACTCTTCATCATTCAGGCCTGCCG 566

     164 gGlyThrGluLeuAspCysGlyIleGluThrAspSerGly 177
```

```
      ||||||||||||||||||||||||||||||||||||||
  567 AGGTACAGAGCTGGACTGCGGTATTGAGACAGACAGTGGA 606

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-890-542A-3

seq_documentation_block:
; Sequence 3, Application US/08890542A
; Patent No. 5840509
;  GENERAL INFORMATION:
;    APPLICANT:  Ni, Binhui
;    APPLICANT:  Paul, Steven M
;    APPLICANT:  Wu, Xin
;    TITLE OF INVENTION:  PROTEASE AND RELATED NUCLEIC ACID
;    TITLE OF INVENTION:  COMPOUNDS
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  United States of America
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/890,542A
;      FILING DATE:  09-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gaylo, Paul J.
;      REGISTRATION NUMBER:  36,808
;      REFERENCE/DOCKET NUMBER:  X-10704
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (317) 276-0756
;      TELEFAX:  (317) 276-3861
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1177 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  mRNA
US-08-890-542A-3


alignment_scores:
            Quality:   30.00              Length:     30
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x US-08-890-542A-3   ..

 Align seg 1/1  to: US-08-890-542A-3  from: 1  to: 1177

     148 CysArgSerLeuThrGlyLysProLysLeuPheIleIleGlnAlaCysAr 164
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     517 UGCCGGAGUCUGACUGGAAAGCCGAAACUCUUCAUCAUUCAGGCCUGCCG 566

     164 gGlyThrGluLeuAspCysGlyIleGluThrAspSerGly 177
         |||||||||||||||||||||||||||||||||||||||
     567 AGGUACAGAGCUGGACUGCGGUAUUGAGACAGACAGUGGA 606

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-800-007A-1

seq_documentation_block:
; Sequence 1, Application US/08800007A
; Patent No. 5834228
;  GENERAL INFORMATION:
;    APPLICANT:  Becker, Joseph
;    APPLICANT:  Nicholson, Donald
;    APPLICANT:  Rotonda, Jennifer
;    APPLICANT:  Thornberry, Nancy
;    APPLICANT:  Fazil, Kimberly
;    APPLICANT:  Gallant, Michel
;    APPLICANT:  Gareau, Yves
;    APPLICANT:  Labelle, Marc
;    APPLICANT:  Peterson, Erin
;    APPLICANT:  Rasper, Dita
;    TITLE OF INVENTION:  CRYSTAL STRUCTURE OF APOPAIN
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  JOSEPH A. COPPOLA - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/800,007A
;      FILING DATE:  13-FEB-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COPPOLA, JOSEPH A
;      REGISTRATION NUMBER:  38,413
;      REFERENCE/DOCKET NUMBER:  19644
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-6734
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  53 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-800-007A-1


alignment_scores:
            Quality:   11.00              Length:     11
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x US-08-800-007A-1   ..

 Align seg 1/1  to: US-08-800-007A-1  from: 1  to: 53

      29 SerGlyIleSerLeuAspAsnSerTyrLysMet 39
         ||||||||||||||||||||||||||||||||||
      20 AGTGGAATATCCCTGGACAACAGTTATAAAATG 52

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-800-007A-5

seq_documentation_block:
; Sequence 5, Application US/08800007A
; Patent No. 5834228
;  GENERAL INFORMATION:
;    APPLICANT:  Becker, Joseph
;    APPLICANT:  Nicholson, Donald
;    APPLICANT:  Rotonda, Jennifer
;    APPLICANT:  Thornberry, Nancy
```

```
;    APPLICANT:  Fazil, Kimberly
;    APPLICANT:  Gallant, Michel
;    APPLICANT:  Gareau, Yves
;    APPLICANT:  Labelle, Marc
;    APPLICANT:  Peterson, Erin
;    APPLICANT:  Rasper, Dita
;    TITLE OF INVENTION:  CRYSTAL STRUCTURE OF APOPAIN
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  JOSEPH A. COPPOLA - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/800,007A
;      FILING DATE:  13-FEB-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COPPOLA, JOSEPH A
;      REGISTRATION NUMBER:  38,413
;      REFERENCE/DOCKET NUMBER:  19644
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-6734
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  53 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-800-007A-5


alignment_scores:
            Quality:   11.00          Length:     11
              Ratio:  1.000            Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x US-08-800-007A-5   ..

 Align seg 1/1  to: US-08-800-007A-5  from: 1  to: 53

      29 SerGlyIleSerLeuAspAsnSerTyrLysMet 39
         |||||||||||||||||||||||||||||||||
      20 AGTGGAATATCCCTGGACAACAGTTATAAAATG 52

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-208-966-24

seq_documentation_block:
; Sequence 24, Application US/09208966
; Patent No. 6221355
; GENERAL INFORMATION:
;  APPLICANT: Dowdy, Steven F.
;  TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
;  FILE REFERENCE: 48881/1742
;  CURRENT APPLICATION NUMBER: US/09/208,966
;  CURRENT FILING DATE: 1998-12-10
;  EARLIER APPLICATION NUMBER: 60/082,402
;  EARLIER FILING DATE: 1998-04-20
;  EARLIER APPLICATION NUMBER: 60/069,012
;  EARLIER FILING DATE: 1997-12-10
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
;   LENGTH: 85
;   TYPE: DNA
;   ORGANISM: human
US-09-208-966-24


alignment_scores:
            Quality:   11.00          Length:     11
              Ratio:  1.000            Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x US-09-208-966-24   ..

 Align seg 1/1  to: US-09-208-966-24  from: 1  to: 85

     164 ArgGlyThrGluLeuAspCysGlyIleGluThr 174
         |||||||||||||||||||||||||||||||||
       6 CGTGGTACCGAACTGGACTGTGGCATTGAGACA 38
```

```
OM of: US-09-613-508-2 to: GenEmbl:*   out_format : pfs

Date: Aug 15, 2002 10:14 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO_spool/US09613508/runat_14082002_112714_7772/app_query.fasta_1.701
 -DB=GenEmbl -QFMT=fastap -SUFFIX=olig.rge -GAPOP=4.500
 -GAPEXT=0.050 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000 -YGAPEXT=60.000
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=oligo
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality
 -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext
 -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US09613508_@CGN1_1_4651 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-613-508-2
Query length: 303
Database: GenEmbl:*
Database sequences: 1797656
Database length: 1873333701
Search time (sec): 2733.790000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000


score_list:
Sequence            Strd Orig    ZScore    EScore Len   ! Documentation  ..
gb_pr:HSU39613        +  254.00 5119.72  8.3e-277  912   ! U39613 Human cysteine protease
gb_pat:AX134417       +  193.00 3876.73  1.4e-207 2309   ! AX134417 Sequence 13 from Pate
gb_pr:HSU37448        +  193.00 3876.73  1.4e-207 2309   ! U37448 Human Mch3 isoform alph
gb_pr:HSU67320        +  193.00 3876.71  1.4e-207 2314   ! U67320 Human Lice2 gamma cyste
gb_pr:HSU40281        +  193.00 3876.65  1.4e-207 2335   ! U40281 Human cysteine protease
gb_pr:HSU67319        +  193.00 3875.89  1.6e-207 2602   ! U67319 Human Lice2 beta cystei
gb_pr:HSU67206        +  193.00 3875.74  1.6e-207 2661   ! U67206 Homo sapiens Lice2 alph
gb_pat:AR103058       +  189.00 3799.30  2.9e-203 1371   ! AR103058 Sequence 1 from paten
gb_pr:BC015799        +  189.00 3795.49  4.8e-203 2361   ! BC015799 Homo sapiens, caspase
gb_pr:HSU37449        +  148.00 2964.36  9.4e-157 2378   ! U37449 Human Mch3 isoform beta
gb_htg:AL627395       +   59.00 1131.04  1.2e-54  155904  ! AL627395 Homo sapiens chromo
gb_pr:AL592546        -   59.00 1130.24  1.4e-54  174776  ! AL592546 Human DNA sequence
gb_ro:D86353          +   44.00  862.57  1.1e-39  964   ! D86353 Mouse mRNA for Mch3, com
gb_ro:AF072124        +   44.00  861.88  1.2e-39 1063   ! AF072124 Rattus norvegicus cas
gb_ro:MAU47332        +   44.00  857.68  2.1e-39 1939   ! U47332 Mesocricetus auratus MC
gb_ro:MMU67321        +   44.00  857.44  2.1e-39 2006   ! U67321 Mus musculus Lice2 cyst
gb_ro:BC005428        +   44.00  856.47  2.4e-39 2305   ! BC005428 Mus musculus, caspase
gb_ro:MMCASP7         +   44.00  856.27  2.5e-39 2372   ! Y13088 M.musculus mRNA for cas
gb_ov:AB038170        +   17.00  312.02  5.1e-09 1534   ! AB038170 Xenopus laevis mRNA f
gb_ov:AB047003        +   16.00  295.89  4.1e-08  849   ! AB047003 Danio rerio mRNA for c
gb_ov:AF083029        +   16.00  292.08  6.6e-08 1464   ! AF083029 Gallus gallus caspase
gb_ov:D89784          +   13.00  234.81  0.0001   882   ! D89784 Xenopus laevis mRNA for
gb_in:AF130469        +   12.00  213.02  0.0017  1096   ! AF130469 Drosophila melanogast
gb_in:AF222007        +   12.00  208.79  0.0029  2006   ! AF222007 Drosophila melanogast
gb_htg:AC008293       +   12.00  185.53  0.0569  55741  ! AC008293 Drosophila melanogas
gb_htg:AC008294       +   12.00  185.43  0.0576  56559  ! AC008294 Drosophila melanogas
gb_htg:AC020380       -   12.00  181.35  0.0973  101381  ! AC020380 Drosophila melanoga
gb_in:AC011615        -   12.00  176.61  0.1787  199653  ! AC011615 Drosophila melanoga
gb_in:AE003713        -   12.00  175.75  0.1995  225795  ! AE003713 Drosophila melanoga
gb_om:AB029345        +   11.00  194.66  0.0176  834   ! AB029345 Sus scrofa mRNA for ca
gb_pr:AF111344        +   11.00  190.66  0.0295 1477   ! AF111344 Homo sapiens caspase-
gb_pr:AF111345        +   11.00  190.24  0.0311 1569   ! AF111345 Homo sapiens caspase-
gb_pat:AR168085       +   11.00  189.68  0.0334 1700   ! AR168085 Sequence 1 from paten
gb_pr:AB038973S5      +   11.00  188.81  0.0373 1924   ! AB038977 Homo sapiens gene for
gb_pr:HSU86214        +   11.00  188.25  0.0401 2085   ! U86214 Human Fas-associated de
gb_pat:BD007034       +   11.00  188.23  0.0403 2092   ! BD007034 Interleukin-1beta con
gb_pat:AR067853       +   11.00  184.56  0.0645 3535   ! AR067853 Sequence 1 from paten
gb_pat:AX134423       +   11.00  184.55  0.0645 3536   ! AX134423 Sequence 19 from Pate
gb_pr:HSU60519        +   11.00  184.55  0.0645 3536   ! U60519 Human apoptotic cys
gb_pr:AC007283        -   11.00  159.48  1.61   127361  ! AC007283 Homo sapiens BA
gb_htg:AC074016       +   11.00  157.09  2.18   179216  ! AC074016 Homo sapiens ch
gb_pat:AX207739       -   10.00  194.36  0.0183  48   ! AX207739 Sequence 37 from Pa
gb_pat:AX207745       -   10.00  194.36  0.0183  48   ! AX207745 Sequence 43 from Pa
gb_sts:G73183         +   10.00  181.00  0.1017  324   ! G73183 csnpcasp9-pcr6-1 Hum
gb_ro:RNU34685        +   10.00  176.25  0.1870  639   ! U34685 Rattus norvegicus cy

seq_name: gb_pr:HSU39613

seq_documentation_block:
LOCUS       HSU39613                912 bp    mRNA    linear   PRI 19-JAN-1996
DEFINITION  Human cysteine protease ICE-LAP3 mRNA, complete cds.
ACCESSION   U39613
VERSION     U39613.1  GI:1125072
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 912)
  AUTHORS   Duan,H., Chinnaiyan,A.M., Hudson,P.L., Wing,J.P., He,W.W. and
            Dixit,V.M.
  TITLE     ICE-LAP3, a novel mammalian homologue of the Caenorhabditis elegans
            cell death protein Ced-3 is activated during Fas- and tumor
            necrosis factor-induced apoptosis
  JOURNAL   J. Biol. Chem. 271 (3), 1621-1625 (1996)
  MEDLINE   96139498
REFERENCE   2  (bases 1 to 912)
  AUTHORS   Duan,H., Chinnaiyan,A.M., Hudson,P.L., Wing,J.P., He,W.W. and
            Dixit,V.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-OCT-1995) Hangjun Duan, Pathology, University of
            Michigan, 1150 W Medical Center Dr MSRBI Room 7514, Ann Arbor, MI
            48109, USA
FEATURES             Location/Qualifiers
     source          1. .912
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .912
                     /function="cysteine protease"
                     /note="member of ICE/CED-3 gene family of cysteine
                     proteases"
                     /codon_start=1
                     /product="ICE-LAP3"
                     /protein_id="AAC50346.1"
                     /db_xref="GI:1125073"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDAIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT       263 a    219 c    223 g    207 t
ORIGIN


alignment_scores:
            Quality:  254.00              Length:    254
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x HSU39613   ..

 Align seg 1/1  to: HSU39613  from: 1  to: 912

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 50

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     51 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 100

     34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    101 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 150

     51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    151 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 200

     67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    201 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 250

     84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    251 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 300

    101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    301 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 350

    117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    351 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 400

    134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    401 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 450

    151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    451 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 500

    167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    501 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 550

    184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    551 AGGCTTGCCGAGGGACCGAGCTTGATGATGCCATCCAGGCCGACTCGGGG 600

    201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    601 CCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGC 650

    217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    651 TGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGA 700

    234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    701 GCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAG 750

    251 GluHisGlyLys 254
        ||||||||||||
    751 GAGCACGGAAAA 762

seq_name: gb_pat:AX134417

seq_documentation_block:
LOCUS       AX134417               2309 bp    DNA     linear   PAT 29-MAY-2001
DEFINITION  Sequence 13 from Patent WO0129232.
ACCESSION   AX134417
VERSION     AX134417.1  GI:14271063
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2309)
  AUTHORS   Cordell,B. and Li,Y.
  TITLE     Functional cloning of genes encoding proteins/enzymes involved in
            proteolytic cleavage
  JOURNAL   Patent: WO 0129232-A 13 26-APR-2001;
            Scios Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2309
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             44. .955
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC39525.1"
                     /db_xref="GI:14271064"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      680 a    453 c    531 g    645 t
ORIGIN


alignment_scores:
           Quality:  193.00              Length:    193
             Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AX134417   ..

 Align seg 1/1  to: AX134417  from: 1  to: 2309

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
      44 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 93

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         |||||||||||||||||||||||||||||||||||||||||||||||||||
      94 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 143

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     144 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 193

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     194 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 243

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     244 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 293

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     294 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 343

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     344 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 393

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     394 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 443

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     444 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 493

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     494 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 543

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||
   544 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 593

   184 lnAlaCysArgGlyThrGluLeuAspAsp 193
       |||||||||||||||||||||||||||||
   594 AGGCTTGCCGAGGGACCGAGCTTGATGAT 622

seq_name: gb_pr:HSU37448

seq_documentation_block:
LOCUS       HSU37448               2309 bp    mRNA    linear   PRI 13-DEC-1995
DEFINITION  Human Mch3 isoform alpha (Mch3) mRNA, complete cds.
ACCESSION   U37448
VERSION     U37448.1  GI:1117846
KEYWORDS    Mch3 isoform alpha.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2309)
  AUTHORS   Fernandes-Alnemri,T., Litwack,G. and Alnemri,E.S.
  TITLE     CPP32, a novel human apoptotic protein with homology to
            Caenorhabditis elegans cell death protein Ced-3 and mammalian
            interleukin-1 beta-converting enzyme
  JOURNAL   J. Biol. Chem. 269 (49), 30761-30764 (1994)
  MEDLINE   95074098
REFERENCE   2  (bases 1 to 2309)
  AUTHORS   Fernandes-Alnemri,T., Takahashi,A., Armstrong,R., Krebs,J.,
            Fritz,L., Tomaselli,K.J., Wang,L., Yu,Z., Croce,C.M., Salveson,G.,
            Earnshaw,W.C., Litwack,G. and Alnemri,E.S.
  TITLE     Mch3, a novel human apoptotic cysteine protease highly related to
            CPP32
  JOURNAL   Cancer Res. 55 (24), 6045-6052 (1995)
  MEDLINE   96105019
REFERENCE   3  (bases 1 to 2309)
  AUTHORS   Alnemri,E.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-1995) Emad S. Alnemri, Pharmacology, Thomas
            Jefferson University, Jefferson Cancer Institute, 233, S. Tenth
            Street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .2309
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="T50-8A.1"
                     /cell_line="Jurkat"
                     /cell_type="T-lymphocyte"
     gene            1. .2309
                     /gene="Mch3"
     CDS             44. .955
                     /gene="Mch3"
                     /function="cysteine protease"
                     /note="new member of ICE-like cysteine protease family"
                     /codon_start=1
                     /evidence=experimental
                     /product="Mch3 isoform alpha"
                     /protein_id="AAC50303.1"
                     /db_xref="GI:1117847"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
     polyA_site      2309
                     /gene="Mch3"
BASE COUNT      680 a    453 c    531 g    645 t
ORIGIN


alignment_scores:
           Quality:  193.00          Length:    193
             Ratio:  1.000             Gaps:      0

 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x HSU37448   ..

 Align seg 1/1  to: HSU37448  from: 1  to: 2309

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      44 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 93

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      94 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 143

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     144 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 193

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     194 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 243

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     244 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 293

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     294 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 343

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     344 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 393

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     394 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 443

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     444 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 493

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     494 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 543

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     544 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 593

     184 lnAlaCysArgGlyThrGluLeuAspAsp 193
         |||||||||||||||||||||||||||||
     594 AGGCTTGCCGAGGGACCGAGCTTGATGAT 622

seq_name: gb_pr:HSU67320

seq_documentation_block:
LOCUS       HSU67320               2314 bp    mRNA    linear   PRI 19-MAR-1997
DEFINITION  Human Lice2 gamma cysteine protease mRNA, complete cds.
ACCESSION   U67320
VERSION     U67320.1  GI:1894914
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2314)
  AUTHORS   Juan,T.S., McNiece,I.K., Argento,J.M., Jenkins,N.A., Gilbert,D.J.,
            Copeland,N.G. and Fletcher,F.A.
  TITLE     Identification and mapping of Casp7, a cysteine protease resembling
            CPP32 beta, interleukin-1 beta converting enzyme, and CED-3
  JOURNAL   Genomics 40 (1), 86-93 (1997)
```

```
  MEDLINE   97224489
REFERENCE   2  (bases 1 to 2314)
  AUTHORS   Juan,T.S.C., McNiece,I.K., Argento,J.M., Jenkins,N.A.,
            Gilbert,D.J., Copeland,N.G. and Fletcher,F.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-1996) Developmental Hematology, Amgen, Inc., 1840
            DeHavilland Dr., Thousand Oaks, CA 91320, USA
FEATURES             Location/Qualifiers
     source          1. .2314
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             40. .951
                     /note="similar to ICE and CED-3 cysteine protease"
                     /codon_start=1
                     /product="Lice2 gamma cysteine protease"
                     /protein_id="AAC51153.1"
                     /db_xref="GI:1894915"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      680 a    459 c    526 g    648 t      1 others
ORIGIN


alignment_scores:
            Quality:  193.00            Length:    193
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x HSU67320   ..

 Align seg 1/1  to: HSU67320  from: 1  to: 2314

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      40 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 89

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      90 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 139

      34 erLeuPheSerLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||
     140 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 189

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     190 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 239

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     240 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 289

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     290 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 339

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     340 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 389

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     390 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 439

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     440 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 489

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     490 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 539

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     540 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 589

     184 lnAlaCysArgGlyThrGluLeuAspAsp 193
         |||||||||||||||||||||||||||||
     590 AGGCTTGCCGAGGGACCGAGCTTGATGAT 618

seq_name: gb_pr:HSU40281

seq_documentation_block:
LOCUS       HSU40281                2335 bp    mRNA    linear   PRI 27-JAN-1996
DEFINITION  Human cysteine protease CMH-1 mRNA, complete cds.
ACCESSION   U40281
VERSION     U40281.1  GI:1167554
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2335)
  AUTHORS   Lippke,J.A., Gu,Y., Sarnecki,C., Caron,P.R. and Su,M.S.
  TITLE     Identification and characterization of CPP32/Mch2 homolog 1, a
            novel cysteine protease similar to CPP32
  JOURNAL   J. Biol. Chem. 271 (4), 1825-1828 (1996)
  MEDLINE   96147144
REFERENCE   2  (bases 1 to 2335)
  AUTHORS   Caron,P.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-NOV-1995) Paul R. Caron, Vertex Pharmaceuticals, 40
            Allston St., Cambridge, MA 02139, USA
FEATURES             Location/Qualifiers
     source          1. .2335
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="IMAGE consortium clone 72778"
     CDS             48. .959
                     /function="cysteine protease"
                     /note="similar to CPP32 and to MCH2; related to Ced3 and
                     interleukin-1 beta converting enzyme"
                     /codon_start=1
                     /product="CMH-1"
                     /protein_id="AAC50352.1"
                     /db_xref="GI:1167555"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      700 a    454 c    535 g    646 t
ORIGIN


alignment_scores:
            Quality:  193.00            Length:    193
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x HSU40281   ..

 Align seg 1/1  to: HSU40281  from: 1  to: 2335

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      48 ATGGCAGATGATCAGGGCTGTATTGAGGAGCAGGGGGTTGAGGATTCAGC 97

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
```

```
      98 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 147

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     148 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 197

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     198 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 247

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     248 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 297

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     298 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 347

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     348 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 397

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     398 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 447

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     448 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 497

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     498 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 547

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     548 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 597

     184 lnAlaCysArgGlyThrGluLeuAspAsp 193
         |||||||||||||||||||||||||||||
     598 AGGCTTGCCGAGGGACCGAGCTTGATGAT 626

seq_name: gb_pr:HSU67319

seq_documentation_block:
LOCUS       HSU67319                2602 bp    mRNA    linear   PRI 19-MAR-1997
DEFINITION  Human Lice2 beta cysteine protease mRNA, complete cds.
ACCESSION   U67319
VERSION     U67319.1  GI:1894912
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2602)
  AUTHORS   Juan,T.S., McNiece,I.K., Argento,J.M., Jenkins,N.A., Gilbert,D.J.,
            Copeland,N.G. and Fletcher,F.A.
  TITLE     Identification and mapping of Casp7, a cysteine protease resembling
            CPP32 beta, interleukin-1 beta converting enzyme, and CED-3
  JOURNAL   Genomics 40 (1), 86-93 (1997)
  MEDLINE   97224489
REFERENCE   2  (bases 1 to 2602)
  AUTHORS   Juan,T.S.C., McNiece,I.K., Argento,J.M., Jenkins,N.A.,
            Gilbert,D.J., Copeland,N.G. and Fletcher,F.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-1996) Developmental Hematology, Amgen, Inc., 1840
            DeHavilland Dr., Thousand Oaks, CA 91320, USA
FEATURES             Location/Qualifiers
     source          1. .2602
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             229. .1239
                     /note="similar to ICE and CED-3 cysteine protease"
                     /codon_start=1
                     /product="Lice2 beta cysteine protease"
                     /protein_id="AAC51152.1"
                     /db_xref="GI:1894913"
                     /translation="MDCVGWPPGRKWHLEKNTSCGGSSGICASYVTQMADDQGCIEEQ
                     GVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVTMRSIKTTRDRVPTYQYNMNFEKLGK
                     CIIINNKNFDKVTGMGVRNGTDKDAEALFKCFRSLGFDVIVYNDCSCAKMQDLLKKAS
                     EEDHTNAACFACILLSHGEENVIYGKDGVTPIKDLTAHFRGDRCKTLLEKPKLFFIQA
                     CRGTELDDGIQADSGPINDTDANPRYKIPVEADFLFAYSTVPGYYSWRSPGRGSWFVQ
                     ALCSILEEHGKDLEIMQILTRVNDRVARHFESQSDDPHFHEKKQIPCVVSMLTKELYF
                     SQ"
BASE COUNT      768 a    515 c    616 g    702 t      1 others
ORIGIN


alignment_scores:
           Quality:  193.00              Length:    193
             Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x HSU67319   ..

 Align seg 1/1  to: HSU67319  from: 1  to: 2602

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     328 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 377

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     378 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 427

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     428 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 477

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     478 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 527

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     528 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 577

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     578 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 627

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     628 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 677

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     678 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 727

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     728 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 777

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     778 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 827

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     828 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 877

     184 lnAlaCysArgGlyThrGluLeuAspAsp 193
         |||||||||||||||||||||||||||||
     878 AGGCTTGCCGAGGGACCGAGCTTGATGAT 906
```

```
seq_name: gb_pr:HSU67206

seq_documentation_block:
LOCUS       HSU67206                2661 bp    mRNA    linear   PRI 03-JAN-2000
DEFINITION  Homo sapiens Lice2 alpha mRNA, complete cds.
ACCESSION   U67206
VERSION     U67206.1  GI:6649572
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2661)
  AUTHORS   Juan,T.S., McNiece,I.K., Argento,J.M., Jenkins,N.A., Gilbert,D.J.,
            Copeland,N.G. and Fletcher,F.A.
  TITLE     Identification and mapping of Casp7, a cysteine protease resembling
            CPP32 beta, interleukin-1 beta converting enzyme, and CED-3
  JOURNAL   Genomics 40 (1), 86-93 (1997)
  MEDLINE   97224489
   PUBMED   9070923
REFERENCE   2  (bases 1 to 2661)
  AUTHORS   Juan,T.S.C., McNiece,I.K., Argento,J.M., Jenkins,N.A.,
            Gilbert,D.J., Copeland,N.G. and Fletcher,F.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-AUG-1996) Developmental Hematology, Amgen, Inc., 1840
            DeHavilland Dr., Thousand Oaks, CA 91320, USA
FEATURES             Location/Qualifiers
     source          1. .2661
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             387. .1298
                     /note="like-ICE cysteine protease"
                     /codon_start=1
                     /product="Lice2 alpha"
                     /protein_id="AAF21460.1"
                     /db_xref="GI:6649573"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      740 a    569 c    642 g    709 t      1 others
ORIGIN


alignment_scores:
            Quality:  193.00             Length:    193
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x HSU67206   ..

 Align seg 1/1  to: HSU67206  from: 1  to: 2661

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     387 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 436

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     437 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 486

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     487 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 536

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     537 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 586

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     587 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 636

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     637 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 686

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     687 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 736

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     737 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 786

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     787 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 836

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     837 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 886

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     887 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 936

     184 lnAlaCysArgGlyThrGluLeuAspAsp 193
         |||||||||||||||||||||||||||||
     937 AGGCTTGCCGAGGGACCGAGCTTGATGAT 965

seq_name: gb_pat:AR103058

seq_documentation_block:
LOCUS       AR103058                1371 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 1 from patent US 6087150.
ACCESSION   AR103058
VERSION     AR103058.1  GI:12814646
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1371)
  AUTHORS   He,W.W., Rosen,C.A., Hudson,P.L. and Hastings,G.A.
  TITLE     Interleukin-1 .beta. converting enzyme like apoptosis protease-3
            and 4
  JOURNAL   Patent: US 6087150-A 1 11-JUL-2000;
FEATURES             Location/Qualifiers
     source          1. .1371
                     /organism="unknown"
BASE COUNT      355 a    357 c    354 g    305 t
ORIGIN


alignment_scores:
            Quality:  189.00             Length:    189
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AR103058   ..

 Align seg 1/1  to: AR103058  from: 1  to: 1371

       5 GlnGlyCysIleGluGluGlnGlyValGluAspSerAlaAsnGluAspSe 21
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     352 CAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTC 401

      21 rValAspAlaLysProAspArgSerSerPheValProSerLeuPheSerL 38
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     402 AGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTA 451
```

```
     38 ysLysLysLysAsnValThrMetArgSerIleLysThrThrArgAspArg 54
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    452 AGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGA 501

     55 ValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGlyLysCysIl 71
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    502 GTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCAT 551

     71 eIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGlyValArgA 88
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    552 CATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAA 601

     88 snGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheArgSerLeu 104
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    602 ACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTG 651

    105 GlyPheAspValIleValTyrAsnAspCysSerCysAlaLysMetGlnAs 121
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    652 GGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGA 701

    121 pLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAlaCysPheA 138
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    702 TCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCG 751

    138 laCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrGlyLysAsp 154
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    752 CCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGAT 801

    155 GlyValThrProIleLysAspLeuThrAlaHisPheArgGlyAspArgCy 171
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    802 GGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATG 851

    171 sLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAlaCysArgG 188
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    852 CAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAG 901

    188 lyThrGluLeuAspAsp 193
        ||||||||||||||||
    902 GGACCGAGCTTGATGAT 918
```

seq_name: gb_pr:BC015799

seq_documentation_block:

LOCUS BC015799 2361 bp mRNA linear PRI 29-OCT-2001
DEFINITION Homo sapiens, caspase 7, apoptosis-related cysteine protease, clone MGC:9176 IMAGE:3921455, mRNA, complete cds.
ACCESSION BC015799
VERSION BC015799.1 GI:16041818
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2361)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (09-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 15 Row: n Column: 12
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 15718696.

FEATURES Location/Qualifiers
source 1. .2361
/organism="Homo sapiens"
/db_xref="LocusID:840"
/db_xref="taxon:9606"
/clone="MGC:9176 IMAGE:3921455"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NIH_MGC_72"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
CDS 79. .990
/codon_start=1
/product="caspase 7, apoptosis-related cysteine protease"
/protein_id="AAH15799.1"
/db_xref="GI:16041819"
/translation="MADEQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT 704 a 464 c 545 g 648 t
ORIGIN

alignment_scores:
Quality: 189.00 Length: 189
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-613-508-2 x BC015799 ..

Align seg 1/1 to: BC015799 from: 1 to: 2361

```
      5 GlnGlyCysIleGluGluGlnGlyValGluAspSerAlaAsnGluAspSe 21
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     91 CAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTC 140

     21 rValAspAlaLysProAspArgSerSerPheValProSerLeuPheSerL 38
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    141 AGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTA 190

     38 ysLysLysLysAsnValThrMetArgSerIleLysThrThrArgAspArg 54
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    191 AGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGA 240

     55 ValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGlyLysCysIl 71
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    241 GTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCAT 290

     71 eIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGlyValArgA 88
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    291 CATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAA 340

     88 snGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheArgSerLeu 104
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    341 ACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTG 390

    105 GlyPheAspValIleValTyrAsnAspCysSerCysAlaLysMetGlnAs 121
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    391 GGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGA 440
```

```
     121 pLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAlaCysPheA 138
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     441 TCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCG 490

     138 laCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrGlyLysAsp 154
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     491 CCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGAT 540

     155 GlyValThrProIleLysAspLeuThrAlaHisPheArgGlyAspArgCy 171
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     541 GGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATG 590

     171 sLysThrLeuGluLysProLysLeuPhePheIleGlnAlaCysArgG 188
         ||||||||||||||||||||||||||||||||||||||||||||||||
     591 CAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAG 640

     188 lyThrGluLeuAspAsp 193
         |||||||||||||||||
     641 GGACCGAGCTTGATGAT 657

seq_name: gb_pr:HSU37449

seq_documentation_block:
LOCUS       HSU37449                2378 bp    mRNA    linear   PRI 13-DEC-1995
DEFINITION  Human Mch3 isoform beta (Mch3) mRNA, complete cds.
ACCESSION   U37449
VERSION     U37449.1  GI:1117848
KEYWORDS    Mch3 isoform beta.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2378)
  AUTHORS   Fernandes-Alnemri,T., Litwack,G. and Alnemri,E.S.
  TITLE     CPP32, a novel human apoptotic protein with homology to
            Caenorhabditis elegans cell death protein Ced-3 and mammalian
            interleukin-1 beta-converting enzyme
  JOURNAL   J. Biol. Chem. 269 (49), 30761-30764 (1994)
  MEDLINE   95074098
REFERENCE   2  (bases 1 to 2378)
  AUTHORS   Fernandes-Alnemri,T., Takahashi,A., Armstrong,R., Krebs,J.,
            Fritz,L., Tomaselli,K.J., Wang,L., Yu,Z., Croce,C.M., Salveson,G.,
            Earnshaw,W.C., Litwack,G. and Alnemri,E.S.
  TITLE     Mch3, a novel human apoptotic cysteine protease highly related to
            CPP32
  JOURNAL   Cancer Res. 55 (24), 6045-6052 (1995)
  MEDLINE   96105019
REFERENCE   3  (bases 1 to 2378)
  AUTHORS   Alnemri,E.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-1995) Emad S Alnemri, Pharmacology, Thomas
            Jefferson University, Jefferson Cancer Institute, 233, S. Tenth
            Street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .2378
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="T50-6A.1"
                     /cell_line="Jurkat"
                     /cell_type="T-lymphocyte"
     gene            1. .2378
                     /gene="Mch3"
     CDS             147. .908
                     /gene="Mch3"
                     /note="does not have protease activity"
                     /codon_start=1
                     /evidence=experimental
                     /product="Mch3 isoform beta"
                     /protein_id="AAC50304.1"
                     /db_xref="GI:1117849"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENMESCSVTQAGVQ
                     RRDLGRLQPPPPRLAEGPSLMMASRPTRGPSMTQMLILDTRSQWKLTSSSPIPRFQAI
                     TRGGAQEEAPGLCKPSAPSWRSTEKTWKSCRSSPG"
     polyA_site      2378
                     /gene="Mch3"
BASE COUNT      680 a    485 c    566 g    647 t
ORIGIN


alignment_scores:
            Quality:  148.00              Length:    148
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x HSU37449   ..

 Align seg 1/1  to: HSU37449  from: 1  to: 2378

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     147 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 196

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     197 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 246

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     247 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 296

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     297 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 346

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     347 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 396

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     397 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 446

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     447 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 496

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     497 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 546

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsn 148
         |||||||||||||||||||||||||||||||||||||||||||||
     547 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAAT 590

seq_name: gb_htg:AL627395

seq_documentation_block:
LOCUS       AL627395               155904 bp    DNA     linear   HTG 08-DEC-2001
DEFINITION  Homo sapiens chromosome 10 clone RP11-575E4, *** SEQUENCING IN
            PROGRESS ***, 3 unordered pieces.
ACCESSION   AL627395
VERSION     AL627395.11  GI:17426548
KEYWORDS    HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Lawlor,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-DEC-2001) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
```

```
COMMENT     On Dec 8, 2001 this sequence version replaced gi:17154386.
            ------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            ------------- Project Information
            Center project name: bA575E4
            ------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 154012 bases at least Q40
            Consensus quality: 154519 bases at least Q30
            Consensus quality: 154817 bases at least Q20
            Insert size: 155704; sum-of-contigs
            Insert size: 160371; 4.2% error; agarose-fp
            Quality coverage: 7.46x in Q20 bases; sum-of-contigs Quality
            coverage: 7.24x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1    81542: contig of 81542 bp in length
            *    81543 81642: gap of      100 bp
            *    81643    89185: contig of 7543 bp in length
            *    89186 89285: gap of      100 bp
            *    89286   155904: contig of 66619 bp in length.
FEATURES             Location/Qualifiers
     source          1. .155904
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-575E4"
                     /clone_lib="RPCI-11.2"
     misc_feature    1. .81542
                     /note="assembly_fragment:03828
                     fragment_chain:1"
     misc_feature    81643. .89185
                     /note="assembly_fragment:03912
                     fragment_chain:1"
     misc_feature    89286. .155904
                     /note="assembly_fragment:02100
                     fragment_chain:1"
BASE COUNT    44589 a  33643 c  33277 g  44195 t    200 others
ORIGIN


alignment_scores:
            Quality:   59.00             Length:     59
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AL627395   ..

 Align seg 1/1  to: AL627395  from: 1  to: 155904

     126 AlaSerGluGluAspHisThrAsnAlaAlaCysPheAlaCysIleLeuLe 142
         |||||||||||||||||||||||||||||||||||||||||||||||||||
  127236 GCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTT 127285

     142 uSerHisGlyGluGluAsnValIleTyrGlyLysAspGlyValThrProI 159
         |||||||||||||||||||||||||||||||||||||||||||||||||||
  127286 AAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAA 127335

     159 leLysAspLeuThrAlaHisPheArgGlyAspArgCysLysThrLeuLeu 175
         |||||||||||||||||||||||||||||||||||||||||||||||||||
  127336 TAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTA 127385

     176 GluLysProLysLeuPhePheIleGln 184
         |||||||||||||||||||||||||||
  127386 GAGAAACCCAAACTCTTCTTCATTCAG 127412

seq_name: gb_pr:AL592546

seq_documentation_block:
LOCUS       AL592546              174776 bp    DNA     linear   PRI 15-NOV-2001
DEFINITION  Human DNA sequence from clone RP11-211N11 on chromosome 10,
            complete sequence.
ACCESSION   AL592546
VERSION     AL592546.7  GI:15591647
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 174776)
  AUTHORS   Bird,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Sep 12, 2001 this sequence version replaced gi:15041956.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 10, constructed by the Sanger Centre Chromosome 10
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr10
            RP11-211N11 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            This sequence is the entire insert of clone RP11-211N11.
FEATURES             Location/Qualifiers
     source          1. .174776
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-211N11"
                     /clone_lib="RPCI-11.1"
BASE COUNT    51535 a  35292 c  35407 g  52542 t
ORIGIN


alignment_scores:
            Quality:   59.00             Length:     59
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AL592546/rev   ..

 Align seg 1/1  to reverse of: AL592546  from: 1  to: 174776
```

```
     126 AlaSerGluGluAspHisThrAsnAlaAlaCysPheAlaCysIleLeuLe 142
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  149222 GCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTT 149173

     142 uSerHisGlyGluGluAsnValIleTyrGlyLysAspGlyValThrProI 159
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  149172 AAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAA 149123

     159 leLysAspLeuThrAlaHisPheArgGlyAspArgCysLysThrLeuLeu 175
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  149122 TAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTA 149073

     176 GluLysProLysLeuPhePheIleGln 184
         |||||||||||||||||||||||||||
  149072 GAGAAACCCAAACTCTTCTTCATTCAG 149046

seq_name: gb_ro:D86353

seq_documentation_block:
LOCUS       D86353                964 bp    mRNA    linear   ROD 07-FEB-1999
DEFINITION  Mouse mRNA for Mch3, complete cds.
ACCESSION   D86353
VERSION     D86353.1  GI:1945545
KEYWORDS    Mch3.
SOURCE      Mus musculus embryonal carcinoma cell cell_line:P19EC cell cDNA to
            mRNA, clone_lib:lambda cDNA library of P19ECcell.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 964)
  AUTHORS   Takashi,M., Koko,U., Takeshi,M. and Mariko,M.Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-1996) Takeshi Mukasa, National Institute of
            Neuroscience,NCNP, Division of Development and Differentiation;
            Ogawahihashi-machi 4-1-1, Kodaira, Tokyo 187, Japan
            (E-mail:momoi@ncnpj.ncnp.go.jp, Tel:81-423-42-2711,
            Fax:81-423-46-1754)
REFERENCE   2  (sites)
  AUTHORS   Mukasa,T., Urase,K., Momoi,M.Y., Kimura,I. and Momoi,T.
  TITLE     Specific expression of CPP32 in sensory neurons of mouse embryos
            and activation of CPP32 in the apoptosis induced by a withdrawal of
            NGF
  JOURNAL   Biochem. Biophys. Res. Commun. 231 (3), 770-774 (1997)
  MEDLINE   97224429
REFERENCE   3  (sites)
  AUTHORS   Mukasa,T., Khoroku,Y., Tsukahara,T., Momoi,M.Y., Kimura,I. and
            Momoi,T.
  TITLE     Wortmannin enhances CPP32-like activity during neuronal
            differentiation of P19 embryonal carcinoma cells induced by
            retinoic acid
  JOURNAL   Biochem. Biophys. Res. Commun. 232 (1), 192-197 (1997)
  MEDLINE   97236307
FEATURES             Location/Qualifiers
     source          1. .964
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /cell_line="P19EC cell"
                     /cell_type="embryonal carcinoma cell"
                     /clone_lib="lambda cDNA library of P19ECcell"
     CDS             53. .964
                     /note="similar to human Mch3alpha"
                     /codon_start=1
                     /product="Mch3"
                     /protein_id="BAA19730.1"
                     /db_xref="GI:1945546"
                     /translation="MTDDQDCAADWEKVDSSSEDGVDAKPDRSSIISSILLKKKRNAS
                     TGPRQTGRDRVPTYLYRMDFQKMGKCIIINNKNFDKATGMDVRNGTDKDAGALFKCFQ
                     NLGFEVTVHNDCSCAKMQDLLRKASEEDHSNSACFACVLLSHGEEDLIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDIDANPRNKIPVEAD
                     FLFAYSTVPGYYSWRNPGKGSWFVQALCSILNEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPRFNEKKQIPCMVSMLTKELYFSR"
BASE COUNT      246 a    260 c    263 g    195 t
ORIGIN


alignment_scores:
            Quality:   44.00             Length:     44
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x D86353   ..

 Align seg 1/1  to: D86353  from: 1  to: 964

     150 IleTyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPh 166
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     500 ATTTACGGGAAAGATGGCGTGACACCCATAAAGGATCTGACAGCTCATTT 549

     166 eArgGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheI 183
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     550 TAGGGGAGACCGATGCAAAACCCTGTTAGAGAAACCCAAACTCTTCTTCA 599

     183 leGlnAlaCysArgGlyThrGluLeuAspAsp 193
         |||||||||||||||||||||||||||||||
     600 TTCAGGCATGCCGAGGGACGGAGCTCGATGAT 631

seq_name: gb_ro:AF072124

seq_documentation_block:
LOCUS       AF072124              1063 bp    mRNA    linear   ROD 24-MAY-2000
DEFINITION  Rattus norvegicus caspase-7 mRNA, complete cds.
ACCESSION   AF072124
VERSION     AF072124.1  GI:3243275
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1063)
  AUTHORS   Forghani,F. and Roy,S.
  TITLE     Rat caspase-7 sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1063)
  AUTHORS   Forghani,F. and Roy,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-1998) Biochemistry, Merck Frosst Centre for
            Therapeutic Research, P.O. Box 1005, Pointe-Claire-Dorval, QUE H9H
            3L1, Canada
FEATURES             Location/Qualifiers
     source          1. .1063
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /tissue_type="spleen"
     CDS             3. .914
                     /codon_start=1
                     /product="caspase-7"
                     /protein_id="AAC24011.1"
                     /db_xref="GI:3243276"
                     /translation="MTDDQDCAAELEMADSSTEDGVDAKPDRSTIISSLLWKKKKNAS
                     MCPVSTTRDRVPTYLYRMDFEKMGKCIIINNKNFDKATGMDVRNGTDKDAEALFKCFR
                     SLGFEVTVYNDCSCAKMQDLLRRASEEDHSNSACFACVLLSHGEENLIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRNPGKGSWFVQALCSILNEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPRFNEKKQIPCMVSMLTKELYFGR"
BASE COUNT      278 a    268 c    275 g    242 t
ORIGIN


alignment_scores:
            Quality:   44.00             Length:     44
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000
```

```
alignment_block:
 US-09-613-508-2 x AF072124   ..

 Align seg 1/1  to: AF072124  from: 1  to: 1063

     150 IleTyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPh 166
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     450 ATTTATGGGAAAGACGGCGTGACACCGATAAAGGATCTGACAGCTCATTT 499

     166 eArgGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheI 183
         |||||||||||||||||||||||||||||||||||||||||||||||||
     500 TAGGGGAGACCGATGCAAAACCCTGCTAGAGAAACCCAAGCTCTTCTTCA 549

     183 leGlnAlaCysArgGlyThrGluLeuAspAsp 193
         |||||||||||||||||||||||||||||||
     550 TCCAGGCGTGCCGAGGGACGGAGCTGGATGAC 581

seq_name: gb_ro:MAU47332

seq_documentation_block:
LOCUS       MAU47332               1939 bp    mRNA    linear   ROD 05-JUN-1996
DEFINITION  Mesocricetus auratus MCH3/SCA-2 mRNA, complete cds.
ACCESSION   U47332
VERSION     U47332.1  GI:1245143
KEYWORDS    .
SOURCE      golden hamster.
  ORGANISM  Mesocricetus auratus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Mesocricetus.
REFERENCE   1  (bases 1 to 1939)
  AUTHORS   Pai,J.T., Brown,M.S. and Goldstein,J.L.
  TITLE     Purification and cDNA cloning of a second apoptosis-related
            cysteine protease that cleaves and activates sterol regulatory
            element binding proteins
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 93 (11), 5437-5442 (1996)
  MEDLINE   96224303
REFERENCE   2  (bases 1 to 1939)
  AUTHORS   Pai,J.-T., Brown,M.S. and Goldstein,J.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-1996) Jih-tung Pai, Molecular Genetics,
            University of Texas, Southwestern Medical Center, 5323 Harry Hines
            Blvd., Dallas, TX 75235-9046, USA
FEATURES             Location/Qualifiers
     source          1. .1939
                     /organism="Mesocricetus auratus"
                     /db_xref="taxon:10036"
     CDS             59. .970
                     /note="cysteine protease"
                     /codon_start=1
                     /product="MCH3/SCA-2"
                     /protein_id="AAC52595.1"
                     /db_xref="GI:1245144"
                     /translation="MADDQNCAPELEKADPSGEDGVDAKPDRSSIISSILGKKKKNAS
                     ACPVKTARDRVPTYLYRMDFEKMGKCIIINNKNFDKVTGMDVRNGTDKDAEALFKCFR
                     SLGFDVVVYNDCSCAKMQDLLRKASEEDHSNSACFACVLLSHGEENLIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGVQADSGPINETDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRNPGKGSWFVQALCSILDEHGKDLEIMQILTRVNDRVARHFESQ
                     CDDPCFNEKKQIPCMVSMLTKELYFGR"
BASE COUNT      493 a    464 c    523 g    459 t
ORIGIN


alignment_scores:
            Quality:   44.00             Length:     44
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x MAU47332   ..

 Align seg 1/1  to: MAU47332  from: 1  to: 1939
```

```
    150 IleTyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPh 166
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    506 ATTTACGGGAAAGATGGCGTGACGCCGATAAAGGACCTGACGGCTCATTT 555

    166 eArgGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheI 183
        |||||||||||||||||||||||||||||||||||||||||||||||||
    556 TAGGGGAGATCGGTGCAAAACCCTGTTAGAGAAACCCAAACTCTTCTTCA 605

    183 leGlnAlaCysArgGlyThrGluLeuAspAsp 193
        |||||||||||||||||||||||||||||||
    606 TCCAGGCATGCCGAGGCACGGAGCTCGATGAC 637
```

OM of: US-09-613-508-2 to: N_Geneseq_032802:* out_format : pfs

Date: Aug 15, 2002 10:21 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool/US09613508/runat_14082002_112714_7815/app_query.fasta_1.701
-DB=N_Geneseq_032802 -QFMT=fastap -SUFFIX=olig.rng -GAPOP=4.500
-GAPEXT=0.050 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000 -YGAPEXT=60.000
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality
-THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09613508_@CGN1_1_190 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-613-508-2
Query length: 303
Database: N_Geneseq_032802:*
Database sequences: 1736436
Database length: 858457221
Search time (sec): 265.980000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81223	+	193.00	3541.19	6.7e-18		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT66992	+	193.00	3534.20	1.7e-18		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03914	+	193.00	3534.20	1.7e-1		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT66970	+	193.00	3534.09	1.7e-18		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT15276	+	189.00	3464.27	1.3e-18		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT66993	+	148.00	2703.03	3.3e-14		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAL26799	-	16.00	294.54	4.7e-0		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL26207	+	12.00	198.81	0.0100		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL26206	+	12.00	188.15	0.0393		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAK28295	-	11.00	185.60	0.0545		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAK41355	-	11.00	185.15	0.0577		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAI66510	+	11.00	181.14	0.0965		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81226	+	11.00	177.02	0.1637		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT90097	+	11.00	175.77	0.1922		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV18819	+	11.00	175.75	0.1926		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV26613	+	11.00	173.99	0.2414		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03917	+	11.00	170.25	0.3902		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13077	-	10.00	184.21	0.0651		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV09402	+	10.00	164.68	0.7966		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03918	+	10.00	163.12	0.9732		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13075	+	10.00	162.68	1.03		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81219	+	10.00	162.67	1.03		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81215	+	10.00	162.46	1.06		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81214	+	10.00	162.33	1.08		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV32608	+	10.00	161.30	1.23		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV32615	+	10.00	161.30	1.23		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT33567	+	10.00	160.20	1.42		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV05471	+	10.00	160.20	1.42		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT15277	+	10.00	160.20	1.42		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV13959	+	10.00	160.08	1.44		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV13960	+	10.00	160.08	1.44		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD15656	+	10.00	159.62	1.53		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAH77927	+	10.00	159.62	1.53		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAH25191	+	10.00	159.62	1.53		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAS12629	+	10.00	159.62	1.53		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81225	+	10.00	159.61	1.53		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL03125	+	10.00	159.54	1.54		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAS10510	+	10.00	159.35	1.58		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAZ46657	+	10.00	158.35	1.8(		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV09401	+	10.00	157.87	1.9		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAC64043	+	10.00	154.28	3.0		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAS10503	+	10.00	154.00	3.		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03910	+	10.00	154.00	3.		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAV80061	+	10.00	153.94	3.1		
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL09303	+	10.00	153.00	3.		

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81223

seq_documentation_block:
ID AAX81223 standard; DNA; 913 BP.
XX
AC AAX81223;
XX
DT 10-SEP-1999 (first entry)
XX
DE Nucleotide sequence of caspase-7 (Mch3).
XX
KW Rev-caspase; cysteine protease; zymogen; caspase; cancer; gene therapy;
KW autoimmune disease; caspase-mediated apoptosis; neurodegenerative;
KW tumour cell; myocardial infarction; human; ds.
XX
OS Homo sapiens.
XX
PN WO9935277-A2.
XX
PD 15-JUL-1999.
XX
PF 11-JAN-1999; 99WO-US00632.
XX
PR 09-JAN-1998; 98US-0070987.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Alnemri ES;
XX
DR WPI; 1999-419353/35.
DR P-PSDB; AAY21721.
XX
PT New isolated nucleic acid molecule encoding a rev-caspase - used
PT for screening and identifying inhibitors or enhancers for treating
PT cancer or autoimmune disease
XX
PS Disclosure; Fig 17A-B; 74pp; English.
XX
CC The invention relates to nucleic acid molecules encoding rev-caspases.
CC Rev-caspases are cysteine proteases that specifically cleave proteins
CC after Asp residues and is expressed as a zymogen, in which the small
CC subunit is N-terminal to a large subunit. A gene delivery vehicle
CC comprising a rev-caspase coding sequence is useful for the treatment of
CC cancer, where the gene delivery vehicle is internalised by tumour cells.
CC The gene delivery vehicle can also be used to treat autoimmune diseases.
CC Cells transfected with a rev-caspase expressing vector can be used in
CC identification of inhibitors or enhancers of caspase-mediated apoptosis.
CC In vitro translated rev-caspase can be used to identify an inhibitor or
CC enhancer of caspase processing activity. Caspase inhibitors are useful
CC for treating neurodegenerative diseases as well as for inhibiting
CC apoptosis in the heart following myocardial infarction. Sequences
CC AAX81217 -AAX81226 represent human caspase genes encoding caspase 1-10
CC gene products (AAY21715-Y21724).
XX
SQ Sequence 913 BP; 263 A; 218 C; 224 G; 207 T; 1 other;

alignment_scores:
Quality: 193.00 Length: 193
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-613-508-2 x AAX81223 ..

```
Align seg 1/1  to: AAX81223  from: 1  to: 913

      1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      1 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 50

     17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 100

     34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    101 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 150

     51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    151 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 200

     67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    201 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 250

     84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    251 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 300

    101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    301 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 350

    117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    351 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 400

    134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    401 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 450

    151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    451 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 500

    167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    501 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 550

    184 lnAlaCysArgGlyThrGluLeuAspAsp 193
        |||||||||||||||||||||||||||||
    551 AGGCTTGCCGAGGGACCGAGCTTGATGAT 579

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT66992

seq_documentation_block:
ID   AAT66992 standard; cDNA; 2309 BP.
XX
AC   AAT66992;
XX
DT   05-AUG-1997  (first entry)
XX
DE   Apototic protease Mch3-alpha cDNA.
XX
KW   Mch3-alpha; cysteine protease; apoptosis; AIDS; ischaemia;
KW   neurodegenerative disease; gene therapy; diagnosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             44..955
FT                   /*tag=  a
FT   intron          595..596
FT                   /*tag=  b
FT                   /note= "site of intron in genomic DNA"
XX
PN   WO9718313-A1.
XX
PD   22-MAY-1997.
XX
PF   12-NOV-1996;   96WO-US18118.
XX
PR   13-NOV-1995;   95US-0556627.
XX
PA   (IDUN-) IDUN PHARM INC.
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES,  Armstrong R,  Fernandes-Alnemri T,  Litwack G;
PI   Tomaselli K;
XX
DR   WPI; 1997-289289/26.
DR   P-PSDB; AAW15262.
XX
PT   New gene encoding Mch3, a cysteine protease that regulates apoptosis
PT   - for treating human diseases associated with apoptosis, and
PT   screening for antagonists and agonists of Mch3
XX
PS   Claim 2; Fig 1; 52pp; English.
XX
CC   A cDNA clone (AAT66992) codes for Mch3-alpha (AAW15262), an apoptotic
CC   protein and novel member of the interleukin-1-beta converting
CC   enzyme (ICE) family of cysteine proteases.  It was obtd. by
CC   screening a human database of expressed sequence tags for sequences
CC   showing homology to ICE family members, and use of a partial cDNA
CC   clone with high homology to CPP32 and Ced-3 to screen a Jurkat cDNA
CC   library.  A second cDNA clone, Mch3-beta (AAT66993), contains a
CC   deletion and insertion corresponding to nucleotides 488-592 of the
CC   Mch3-alpha sequence.  Mch3 nucleic acids can be used to produce
CC   recombinant Mch3 polypeptides useful for treating diseases
CC   associated with cell death, such as AIDS, ischaemic injury,
CC   neurodegenerative diseases, etc.  They can also be used in gene
CC   therapy to regulate apoptosis.
XX
SQ   Sequence 2309 BP; 680 A; 453 C; 531 G; 645 T; 0 other;


alignment_scores:
            Quality:  193.00              Length:    193
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AAT66992   ..

 Align seg 1/1  to: AAT66992  from: 1  to: 2309

      1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     44 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 93

     17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     94 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 143

     34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    144 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 193

     51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    194 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 243

     67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    244 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 293

     84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    294 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 343
```

```
     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     344 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 393

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     394 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 443

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     444 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 493

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     494 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 543

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     544 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 593

     184 lnAlaCysArgGlyThrGluLeuAspAsp 193
         |||||||||||||||||||||||||||||
     594 AGGCTTGCCGAGGGACCGAGCTTGATGAT 622

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03914

seq_documentation_block:
ID   AAD03914 standard; DNA; 2309 BP.
XX
AC   AAD03914;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Human caspase-7 DNA.
XX
KW   Human; caspase-7; interleukin-1 converting enzyme; ICE-LAP3; CHM-1; Mhc3;
KW   cysteine protease; apoptosis; caspase expression cassette; metastasis;
KW   tumour; cathepsin B; urokinase; proliferation; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             44..955
FT                   /*tag= a
FT                   /product= "Human caspase-7"
XX
PN   WO200129232-A2.
XX
PD   26-APR-2001.
XX
PF   19-OCT-2000; 2000WO-US28941.
XX
PR   20-OCT-1999;   99US-0160559.
PR   14-AUG-2000; 2000US-0225564.
XX
PA   (SCIO-) SCIOS INC.
XX
PI   Cordell B,  Li Y;
XX
DR   WPI; 2001-290920/30.
DR   P-PSDB; AAE00604.
XX
PT   Novel fusion polypeptide comprising first and second caspase subunit
PT   separated by cleavage site not associated in nature with caspase
PT   subunit, useful for cloning gene encoding enzymes involved in
PT   proteolytic cleavage  -
XX
PS   Disclosure; Fig 13; 116pp; English.
XX
CC   The present sequence is a DNA encoding human Caspase-7 also known as
CC   interleukin-1 converting enzyme (ICE) LAP3, CHM-1 and Mhc3. Caspases are
CC   a family of cysteine proteases, that participate in the initiation and
CC   execution of apoptosis. Caspases exist as pro-enzymes, activated by
CC   cleavage into a large and small subunit, occurring after specific
CC   aspartic acid residues within the pro-enzyme sequence.
CC   The present invention relates to a method for functional cloning of genes
CC   encoding proteins or enzymes involved in proteolytic cleavage. The
CC   invention is based on the use of caspase expression cassettes comprising
CC   the coding sequence of a proteolytic cleavage site flanked by sequences
CC   encoding two caspase subunits. A fusion polypeptide comprising a first
CC   and a second caspase subunit, separated by a cleavage site not associated
CC   in nature, is useful for cloning gene encoding enzymes involved in
CC   proteolytic cleavage. An expression cassette containing fusion
CC   polypeptide is used to identify a mutant cell line deficient in an
CC   enzyme of interest and is also useful for diagnosis and suppression of
CC   proliferation or metastases of a tumour cell characterised by
CC   overexpression of a polypeptide (e.g. Cathepsin B or urokinase,
CC   selectively expressed in the tumour cells). DNA encoding fusion
CC   polypeptide is used in gene therapy.
XX
SQ   Sequence 2309 BP; 680 A; 453 C; 531 G; 645 T; 0 other;


alignment_scores:
            Quality:  193.00              Length:    193
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AAD03914    ..

 Align seg 1/1  to: AAD03914  from: 1  to: 2309

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      44 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 93

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      94 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 143

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     144 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 193

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     194 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 243

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     244 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 293

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     294 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 343

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     344 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 393

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     394 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 443

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     444 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 493

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     494 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 543

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     544 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 593

     184 lnAlaCysArgGlyThrGluLeuAspAsp 193
         ||||||||||||||||||||||||||||||
     594 AGGCTTGCCGAGGGACCGAGCTTGATGAT 622

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT66970

seq_documentation_block:
ID   AAT66970 standard; cDNA; 2341 BP.
XX
AC   AAT66970;
XX
DT   21-JUL-1997  (first entry)
XX
DE   Cysteine protease CMH-1 cDNA.
XX
KW   Cysteine protease; CMH-1; CPP32/Mch2 homologue-1; apoptosis;
KW   programmed cell death; cancer; neurodegenerative disease;
KW   autoimmune disease; gene therapy; diagnosis; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             54..965
FT                   /*tag= a
FT                   /product= CMH-1 proenzyme
FT                   /note= "bases 123-962 and 138-962 encode active
FT                         active CMH-1 (Claim 2)"
XX
PN   WO9716552-A1.
XX
PD   09-MAY-1997.
XX
PF   01-NOV-1996;   96WO-US17431.
XX
PR   16-NOV-1995;   95US-0558733.
PR   03-NOV-1995;   95US-0007211.
PR   06-NOV-1995;   95US-0007251.
XX
PA   (VERT-) VERTEX PHARM INC.
XX
PI   Lippke JA,  Su M;
XX
DR   WPI; 1997-272121/24.
DR   P-PSDB; AAW15247.
XX
PT   DNA encoding active, activatable or inactive cysteine protease CMH-1
PT   - useful in gene therapy for promoting and inhibiting apoptosis and
PT   for diagnosing cells with potential for apoptosis
XX
PS   Claim 1; Page 30-32; 48pp; English.
XX
CC   A cDNA clone (T66970) includes to coding sequence for the proenzyme
CC   (W15247) of CMH-1, a human cysteine protease involved in apoptosis.
CC   To obtain the clone, the GenBank Expressed Sequence Tag (EST)
CC   database was screened for clones encoding polypeptides similar to
CC   interleukin-1b converting enzyme.  Partial cDNA clone T50828 was
CC   identified.  Partial cDNA clone 72778 corresponding to this EST was
CC   then obtd.  A spleen cDNA pool was subjected to PCR using primers
CC   (T66971-92) based on a 5' adaptor and EST T50828.  A PCR fragment
CC   contg. the 5' coding region was ligated to the clone 72778 3' region
CC   to obtain the full-length CHM-1 sequence.  CHM-1 (i.e. CPP32/Mch2
CC   homologue-1) DNA can be used in gene therapy to promote apoptosis,
CC   esp. in cancer cells, and to modulate patient cholesterol levels,
CC   as well as to diagnose cells with potential for apoptosis.
CC   Antisense sequences or sequences encoding inactive CHM-1 mutants
CC   can be used to inhibit apoptosis, e.g. to prevent neurological
CC   disease.
XX
SQ   Sequence 2341 BP; 702 A; 455 C; 536 G; 648 T; 0 other;

alignment_scores:
           Quality:  193.00              Length:    193
             Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AAT66970   ..

 Align seg 1/1  to: AAT66970  from: 1  to: 2341

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      54 ATGGCAGATGATCAGGGCTGTATTGAGGAGCAGGGGGTTGAGGATTCAGC 103

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     104 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 153

      34 erLeuPheSerLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     154 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 203

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     204 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 253

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     254 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 303

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     304 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 353

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     354 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 403

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     404 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 453

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     454 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 503

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     504 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 553

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     554 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 603

     184 lnAlaCysArgGlyThrGluLeuAspAsp 193
         ||||||||||||||||||||||||||||||
     604 AGGCTTGCCGAGGGACCGAGCTTGATGAT 632

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT15276

seq_documentation_block:
ID   AAT15276 standard; cDNA; 1371 BP.
XX
AC   AAT15276;
XX
DT   27-OCT-1996  (first entry)
XX
DE   Human interleukin-I-converting enzyme-like apoptosis protease-3 cDNA.
XX
KW   ICE-LAP-3; interleukin-I-converting enzyme-like apoptosis protease;
KW   enzyme; Alzheimer's disease; Parkinson's disease; septic shock;
KW   head injury; rheumatoid arthritis; ss.
XX
```

```
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             115..1137
FT                   /*tag= a
XX
PN   WO9613603-A1.
XX
PD   09-MAY-1996.
XX
PF   06-JUN-1995;   95WO-US07235.
XX
PR   01-NOV-1994;   94US-0334251.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Hastings GA,  He W,  Hudson PL,  Rosen CA;
XX
DR   WPI; 1996-239509/24.
DR   P-PSDB; AAR95830.
XX
PT   Human ICE-LAP-3 and -4 DNA and protein  - useful in the diagnosis
PT   and treatment of Alzheimer's disease, Parkinson's disease,
PT   rheumatoid arthritis, septic shock and head injury
XX
PS   Claim 1; Page 91-92; 67pp; English.
XX
CC   This sequence was derived from a cDNA library from human endometrial
CC   tumuor and may be used for the recombinant production of the ICE-LAP-
CC   3 protein e.g. in COS cells.  The encoded protein may be used
CC   therapeutically, e.g. as an antitumor or antiviral agent and to
CC   control embryonic development and tissue homeostasis.  The protein
CC   can also be used to treat immunosuppression disorders, such as AIDS,
CC   by targeting virus infected cells for cell death.  The DNA may find
CC   use in gene therapy applications.
XX
SQ   Sequence 1371 BP; 355 A; 357 C; 354 G; 305 T; 0 other;


alignment_scores:
            Quality:  189.00              Length:    189
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AAT15276   ..

 Align seg 1/1  to: AAT15276  from: 1  to: 1371

       5 GlnGlyCysIleGluGluGlnGlyValGluAspSerAlaAsnGluAspSe 21
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     352 CAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTC 401

      21 rValAspAlaLysProAspArgSerSerPheValProSerLeuPheSerL 38
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     402 AGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTA 451

      38 ysLysLysLysAsnValThrMetArgSerIleLysThrThrArgAspArg 54
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     452 AGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGA 501

      55 ValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGlyLysCysIl 71
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     502 GTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCAT 551

      71 eIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGlyValArgA 88
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     552 CATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAA 601

      88 snGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheArgSerLeu 104
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     602 ACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTG 651

     105 GlyPheAspValIleValTyrAsnAspCysSerCysAlaLysMetGlnAs 121
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     652 GGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGA 701

     121 pLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAlaCysPheA 138
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     702 TCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCG 751

     138 laCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrGlyLysAsp 154
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     752 CCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGAT 801

     155 GlyValThrProIleLysAspLeuThrAlaHisPheArgGlyAspArgCy 171
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     802 GGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATG 851

     171 sLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAlaCysArgG 188
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     852 CAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAG 901

     188 lyThrGluLeuAspAsp 193
         |||||||||||||||||
     902 GGACCGAGCTTGATGAT 918

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT66993

seq_documentation_block:
ID   AAT66993 standard; cDNA; 2377 BP.
XX
AC   AAT66993;
XX
DT   05-AUG-1997  (first entry)
XX
DE   Cysteine protease Mch3-beta cDNA.
XX
KW   Mch3-beta; cysteine protease; apoptosis; AIDS; ischaemia;
KW   neurodegenerative disease; gene therapy; diagnosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             146..839
FT                   /*tag= a
XX
PN   WO9718313-A1.
XX
PD   22-MAY-1997.
XX
PF   12-NOV-1996;   96WO-US18118.
XX
PR   13-NOV-1995;   95US-0556627.
XX
PA   (IDUN-) IDUN PHARM INC.
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES,  Armstrong R,  Fernandes-Alnemri T,  Litwack G;
PI   Tomaselli K;
XX
DR   WPI; 1997-289289/26.
DR   P-PSDB; AAW15263.
XX
PT   New gene encoding Mch3, a cysteine protease that regulates apoptosis
PT   - for treating human diseases associated with apoptosis, and
PT   screening for antagonists and agonists of Mch3
XX
PS   Claim 2; Fig 1; 52pp; English.
XX
CC   A cDNA clone (AAT66993) codes for Mch3-beta (AAW15263), a novel member
CC   of the interleukin-1-beta converting enzyme (ICE) family of
CC   cysteine proteases.  It was obtd. by screening a human database of
CC   expressed sequence tags for sequences showing homology to ICE
CC   family members, and use of a partial cDNA clone with high homology
CC   to CPP32 and Ced-3 to screen a Jurkat cDNA library.  Mch3-beta cDNA
```

```
CC   contains a deletion and insertion corresponding to nucleotides
CC   488-592 of Mch3-alpha cDNA (AAT66992) and also has a longer 3'
CC   nontranslated sequence.  Mch3-beta lacks the active site of Mch3-
CC   alpha.  Mch3 nucleic acids can be used to produce recombinant Mch3
CC   polypeptides useful for treating diseases associated with cell
CC   death, such as AIDS, ischaemic injury, neurodegenerative diseases,
CC   etc.  They can also be used in gene therapy to regulate apoptosis.
XX
SQ   Sequence 2377 BP; 680 A; 485 C; 565 G; 647 T; 0 other;


alignment_scores:
            Quality:  148.00              Length:    148
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AAT66993   ..

 Align seg 1/1  to: AAT66993  from: 1  to: 2377

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     146 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 195

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     196 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 245

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     246 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 295

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     296 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 345

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     346 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 395

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     396 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 445

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     446 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 495

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     496 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 545

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsn 148
         ||||||||||||||||||||||||||||||||||||||||||||
     546 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAAT 589

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAL26799

seq_documentation_block:
ID   AAL26799 standard; DNA; 51 BP.
XX
AC   AAL26799;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #7.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 1402; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 11 A; 14 C; 14 G; 12 T; 0 other;


alignment_scores:
            Quality:   16.00              Length:     16
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AAL26799/rev   ..

 Align seg 1/1  to reverse of: AAL26799  from: 1  to: 51

     224 SerThrValProGlyTyrTyrSerTrpArgSerProGlyArgGlySer 239
         ||||||||||||||||||||||||||||||||||||||||||||||||
      49 TCCACGGTTCCAGGCTATTACTCATGGAGGAGCCCAGGAAGAGGCTCC 2

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL26207

seq_documentation_block:
ID   ABL26207 standard; DNA; 927 BP.
XX
AC   ABL26207;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster genomic polynucleotide SEQ ID NO 30094.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
```

```
KW   pharmaceutical; gene; ds.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 30094; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 927 BP; 235 A; 284 C; 242 G; 166 T; 0 other;


alignment_scores:
           Quality:   12.00            Length:     12
             Ratio:  1.000              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x ABL26207    ..

 Align seg 1/1  to: ABL26207  from: 1  to: 927

     177 LysProLysLeuPhePheIleGlnAlaCysArgGly 188
         ||||||||||||||||||||||||||||||||||||
     484 AAACCCAAACTCTTCTTCATCCAGGCCTGTCGCGGA 519

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL26206

seq_documentation_block:
ID   ABL26206 standard; DNA; 3809 BP.
XX
AC   ABL26206;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster genomic polynucleotide SEQ ID NO 30091.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ds.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 30091; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3809 BP; 1023 A; 868 C; 845 G; 1073 T; 0 other;


alignment_scores:
           Quality:   12.00            Length:     12
             Ratio:  1.000              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x ABL26206    ..

 Align seg 1/1  to: ABL26206  from: 1  to: 3809

     177 LysProLysLeuPhePheIleGlnAlaCysArgGly 188
         ||||||||||||||||||||||||||||||||||||
    2061 AAACCCAAACTCTTCTTCATCCAGGCCTGTCGCGGA 2096

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAK28295

seq_documentation_block:
ID   AAK28295 standard; DNA; 462 BP.
XX
AC   AAK28295;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 2852.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00668.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
```

```
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human bone marrow -
XX
PS   Example 4; SEQ ID NO: 2852; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention.
XX
SQ   Sequence 462 BP; 146 A; 102 C; 115 G; 99 T; 0 other;


alignment_scores:
            Quality:   11.00              Length:     11
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AAK28295/rev   ..

 Align seg 1/1  to reverse of: AAK28295  from: 1  to: 462

     176 GluLysProLysLeuPhePheIleGlnAlaCys 186
         ||||||||||||||||||||||||||||||||||
     334 GAAAAACCTAAACTCTTTTTCATCCAGGCCTGC 302

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAK41355

seq_documentation_block:
ID   AAK41355 standard; DNA; 490 BP.
XX
AC   AAK41355;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 15912.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00668.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human bone marrow -
XX
PS   Example 4; SEQ ID NO: 15912; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention.
XX
SQ   Sequence 490 BP; 126 A; 123 C; 128 G; 113 T; 0 other;


alignment_scores:
            Quality:   11.00              Length:     11
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AAK41355/rev   ..

 Align seg 1/1  to reverse of: AAK41355  from: 1  to: 490

     176 GluLysProLysLeuPhePheIleGlnAlaCys 186
         ||||||||||||||||||||||||||||||||||
     242 GAAAAACCTAAACTCTTTTTCATCCAGGCCTGC 210

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAI66510

seq_documentation_block:
ID   AAI66510 standard; DNA; 834 BP.
XX
AC   AAI66510;
XX
DT   11-DEC-2001  (first entry)
XX
DE   Pig caspase coding sequence #2.
XX
KW   Pig; caspase; inflammation; infection; ds.
XX
OS   Sus scrofa.
XX
PN   JP2001169785-A.
XX
PD   26-JUN-2001.
XX
PF   03-OCT-2000; 2000JP-0303931.
XX
PR   05-OCT-1999;   99JP-0284829.
XX
PA   (NORQ ) NORINSUISANSHO KACHIKU EISEI.
PA   (MUNE/) MUNETA Y.
PA   (MORI/) MORI Y.
PA   (SHIM/) SHIMOCHI Y.
PA   (ARAI/) ARAI K.
XX
DR   WPI; 2001-592539/67.
DR   P-PSDB; AAG78712.
XX
PT   Structural gene of swine caspase family, used for artificial control of
PT   inflammatory reactions in infectious diseases -
XX
PS   Example 2; Page 15-16; 23pp; Japanese.
XX
CC   The present invention provides two versions of a porcine caspase protein.
CC   These can be used to control inflammatory reactions to microbial or viral
CC   infectious diseases. The present sequence is the coding sequence of one
CC   version of the protein of the invention.
XX
```

```
SQ   Sequence 834 BP; 276 A; 149 C; 179 G; 230 T; 0 other;


alignment_scores:
            Quality:   11.00              Length:     11
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AAI66510   ..

 Align seg 1/1  to: AAI66510  from: 1  to: 834

      70 CysIleIleIleAsnAsnLysAsnPheAspLys 80
         |||||||||||||||||||||||||||||||||
     139 TGTATAATAATTAATAATAAGAACTTTGATAAA 171

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81226

seq_documentation_block:
ID   AAX81226 standard; DNA; 1440 BP.
XX
AC   AAX81226;
XX
DT   10-SEP-1999  (first entry)
XX
DE   Nucleotide sequence of caspase-10 (Mch4).
XX
KW   Rev-caspase; cysteine protease; zymogen; caspase; cancer; gene therapy;
KW   autoimmune disease; caspase-mediated apoptosis; neurodegenerative;
KW   tumour cell; myocardial infarction; human; ds.
XX
OS   Homo sapiens.
XX
PN   WO9935277-A2.
XX
PD   15-JUL-1999.
XX
PF   11-JAN-1999;   99WO-US00632.
XX
PR   09-JAN-1998;   98US-0070987.
XX
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES;
XX
DR   WPI; 1999-419353/35.
DR   P-PSDB; AAY21724.
XX
PT   New isolated nucleic acid molecule encoding a rev-caspase  - used
PT   for screening and identifying inhibitors or enhancers for treating
PT   cancer or autoimmune disease
XX
PS   Disclosure; Fig 20A-B; 74pp; English.
XX
CC   The invention relates to nucleic acid molecules encoding rev-caspases.
CC   Rev-caspases are cysteine proteases that specifically cleave proteins
CC   after Asp residues and is expressed as a zymogen, in which the small
CC   subunit is N-terminal to a large subunit. A gene delivery vehicle
CC   comprising a rev-caspase coding sequence is useful for the treatment of
CC   cancer, where the gene delivery vehicle is internalised by tumour cells.
CC   The gene delivery vehicle can also be used to treat autoimmune diseases.
CC   Cells transfected with a rev-caspase expressing vector can be used in
CC   identification of inhibitors or enhancers of caspase-mediated apoptosis.
CC   In vitro translated rev-caspase can be used to identify an inhibitor or
CC   enhancer of caspase processing activity. Caspase inhibitors are useful
CC   for treating neurodegenerative diseases as well as for inhibiting
CC   apoptosis in the heart following myocardial infarction. Sequences
CC   AAX81217 -AAX81226 represent human caspase genes encoding caspase 1-10
CC   gene products (AAY21715-Y21724).
XX
SQ   Sequence 1440 BP; 419 A; 343 C; 335 G; 343 T; 0 other;


alignment_scores:
            Quality:   11.00              Length:     11
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AAX81226   ..

 Align seg 1/1  to: AAX81226  from: 1  to: 1440

     176 GluLysProLysLeuPhePheIleGlnAlaCys 186
         |||||||||||||||||||||||||||||||||
    1042 GAAAAACCTAAACTCTTTTTCATCCAGGCCTGC 1074

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT90097

seq_documentation_block:
ID   AAT90097 standard; DNA; 1700 BP.
XX
AC   AAT90097;
XX
DT   09-APR-1998  (first entry)
XX
DE   Mch4 coding sequence.
XX
KW   Mch4; Mch5; aspartic acid specific Cys protease; cell apoptosis; stroke;
KW   increased cell survival; hormone dependent tumour; autoimmune disease;
KW   immunoglobulin mediated glomerulonephritis; degenerative disease;
KW   therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             148..1587
FT                   /*tag= a
XX
PN   WO9735020-A1.
XX
PD   25-SEP-1997.
XX
PF   19-MAR-1997;   97WO-US04330.
XX
PR   14-JUN-1996;   96US-0665220.
PR   19-MAR-1996;   96US-0618408.
XX
PA   (IDUN-) IDUN PHARM INC.
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES,  Armstrong R,  Fernandes-Alnemri T,  Litwack G;
PI   Tomaselli K;
XX
DR   WPI; 1997-480225/44.
DR   P-PSDB; AAW27390.
XX
PT   Aspartic acid specific cysteine protease(s) Mch4 and Mch5  - which
PT   are involved in cell apoptosis, useful to diagnose and treat, e.g.
PT   cancer, autoimmune, Alzheimer's or Parkinson's disease
XX
PS   Claim 2; Fig 1; 76pp; English.
XX
CC   This sequence represents the Mch4 gene, and is an isolated gene of the
CC   invention. Mch4 and Mch5 (see AAW27391) are members of the Aspartic acid
CC   specific Cys protease family involved in cell apoptosis. The genes and
CC   proteins can be to diagnose, treat or reduce the severity of diseases
CC   resulting from increased cell survival, e.g. hormone dependent tumours
CC   such as breast, prostate or ovarian cancers, or autoimmune diseases,
CC   such as systemic lupus erythematosus or immunoglobulin mediated
CC   glomerulonephritis, diseases resulting from decreased cell survival,
CC   e.g. degenerative diseases such as Alzheimer's or Parkinson's disease,
CC   or amyotrophic lateral sclerosis or other diseases associated with
CC   increased apoptosis such as aplastic anaemia, stroke, ischaemic injury
CC   following myocardial infarction or reperfusion injury.
```

```
XX
SQ   Sequence 1700 BP; 484 A; 405 C; 392 G; 417 T; 2 other;


alignment_scores:
            Quality:   11.00              Length:     11
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AAT90097   ..

 Align seg 1/1  to: AAT90097  from: 1  to: 1700

     176 GluLysProLysLeuPhePheIleGlnAlaCys 186
         |||||||||||||||||||||||||||||||||
    1189 GAAAAACCTAAACTCTTTTTCATCCAGGCCTGC 1221

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV18819

seq_documentation_block:
ID   AAV18819 standard; cDNA; 1704 BP.
XX
AC   AAV18819;
XX
DT   11-JUN-1998  (first entry)
XX
DE   Human protease FMH-1 cDNA.
XX
KW   Human; protease FMH-1; interleukin-1 beta converting enzyme;
KW   ICE/CED3-like protease; apoptosis induction;
KW   apoptosis inhibition; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             121..1686
FT                   /*tag= a
FT                   /product= protease_FMH-1
XX
PN   WO9800554-A1.
XX
PD   08-JAN-1998.
XX
PF   02-JUL-1997;   97WO-US11590.
XX
PR   03-JUL-1996;   96US-0675123.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Bowman MR;
XX
DR   WPI; 1998-086979/08.
DR   P-PSDB; AAW50007.
XX
PT   Isolated nucleic acid encoding human FMH-1 protein - useful for
PT   treating conditions associated with failed or excess apoptosis e.g.
PT   cancer and neurological diseases
XX
PS   Claim 1; Pages 15-16; 29pp; English.
XX
CC   The present sequence encodes the human protease FMH-1, which is an
CC   interleukin-1 beta converting enzyme (ICE)/CED3-like protease.
CC   FMH-1 can be used to induce apoptosis, while fragments may
CC   induce or inhibit apoptosis depending on which FMH-1 region they
CC   bind. Peptide mimetics or small molecule inhibitors targetted to
CC   FMH-1 can have the same effects. Induction of apoptosis can be
CC   used, e.g. to treat autoimmune disease, cancer and persistent
CC   infections, while its inhibition is used to treat
CC   neurodegeneration, e.g. Alzheimer's, Parkinson's or Huntington's
CC   diseases, or ischaemic damage, also to increase response to weakly
CC   immunogenic antigens. FMH-1 can also be used to generate
CC   antibodies for therapeutic use by blocking binding of ligands to
CC   FMH-1.
XX
SQ   Sequence 1704 BP; 481 A; 411 C; 413 G; 399 T; 0 other;


alignment_scores:
            Quality:   11.00              Length:     11
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AAV18819   ..

 Align seg 1/1  to: AAV18819  from: 1  to: 1704

     176 GluLysProLysLeuPhePheIleGlnAlaCys 186
         |||||||||||||||||||||||||||||||||
    1291 GAAAAACCTAAACTCTTTTTCATCCAGGCCTGC 1323
```

```
OM of: US-09-613-508-2 to: EST:*   out_format : pfs

Date: Aug 15, 2002  9:29 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=x1h
 -Q=/cgn2_1/USPTO_spool/US09613508/runat_14082002_112713_7761/app_query.fasta_1.701
 -DB=EST -QFMT=fastap -SUFFIX=olig.rst -GAPOP=4.500 -GAPEXT=0.050
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=60.000 -YGAPEXT=60.000 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15
 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500 -MINLEN=0
 -MAXLEN=2000000000 -USER=US09613508_@CGN1_1_3002 -NCPU=6
 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-09-613-508-2
Query length: 303
Database: EST:*
Database sequences: 13736207
Database length: -1841457050
Search time (sec): 2160.120000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000


score_list:
Sequence          Strd Orig    ZScore    EScore Len   ! Documentation  ..
gb_est1:AL542189    +  193.00 3065.15  1.7e-161  973   ! AL542189 AL542189 LTI_FL002_PL1
gb_est1:AU130224    +  183.00 2906.77  1.2e-152  774   ! AU130224 AU130224 NT2RP3 Homo s
gb_est2:BI821936    +  176.00 2794.27  2.1e-146  926   ! BI821936 603039745F1 NIH_MGC_11
gb_est1:AU126220    +  159.00 2523.79  2.5e-131  809   ! AU126220 AU126220 NT2RP1 Homo s
gb_est2:BI824214    +  147.00 2332.80  1.1e-120  745   ! BI824214 603040584F1 NIH_MGC_11
gb_est2:BM150092    +  144.00 2286.57  4.1e-118  532   ! BM150092 TCAAP4Q11752 Pediatric
gb_est2:BM194512    +  142.00 2255.00  2.3e-116  497   ! BM194512 TCAAP2Q11752 Pediatric
gb_est2:BG339524    +  141.00 2235.67  2.8e-115 1006   ! BG339524 602437368F1 NIH_MGC_4
gb_est2:BG025549    +  139.00 2204.70  1.5e-113  830   ! BG025549 602274928F1 NIH_MGC_85
gb_est2:BE619147    +  138.00 2188.95  1.1e-112  796   ! BE619147 601472883F1 NIH_MGC_68
gb_est2:BE891855    +  136.00 2156.15  7.5e-111  961   ! BE891855 601435409F1 NIH_MGC_72
gb_est2:BM147460    +  134.00 2127.53  2.9e-109  485   ! BM147460 TCAAP1Q11752 Pediatric
gb_est2:BF971443    +  131.00 2078.24  1.6e-106  656   ! BF971443 602272969F1 NIH_MGC_84
gb_est2:BM459840    +  123.00 1948.28  2.8e-99  1075   ! BM459840 AGENCOURT_6422147 NIH
gb_est1:AU143139    +  116.00 1838.57  3.7e-93   719   ! AU143139 AU143139 Y79AA1 Homo s
gb_est2:BF540869    +  115.00 1822.06  3.0e-92   808   ! BF540869 602067587F1 NIH_MGC_58
gb_est1:AU131461    +  115.00 1821.87  3.1e-92   840   ! AU131461 AU131461 NT2RP3 Homo s
gb_est1:AA101428    +  105.00 1665.92  1.5e-83   402   ! AA101428 zn72e08.r1 Stratagene
gb_est2:BE894817    +  104.00 1647.89  1.5e-82   621   ! BE894817 601436392F1 NIH_MGC_72
gb_est2:BG674973    +  104.00 1646.02  1.9e-82   917   ! BG674973 602621183F1 NIH_CGAP_S
gb_est1:AU134778    +   92.00 1455.35  8.1e-72   791   ! AU134778 AU134778 PLACE1 Homo s
gb_est2:BF668893    +   70.00 1104.16  3.0e-52   846   ! BF668893 602123296F1 NIH_MGC_56
gb_est2:BI458770    +   66.00 1040.90  9.9e-49   756   ! BI458770 603200030F1 NIH_MGC_97
gb_est2:BE893968    +   60.00  946.08  1.9e-43   630   ! BE893968 601437754F1 NIH_MGC_72
gb_est1:AU117167    +   60.00  945.58  2.0e-43   700   ! AU117167 AU117167 HEMBA1 Homo s
gb_est1:AA298267    +   56.00  885.03  4.8e-40   356   ! AA298267 EST114122 Prostate gla
gb_est1:AA297423    +   51.00  806.27  1.2e-35   290   ! AA297423 EST112974 Endometrial
gb_est2:H91868      +   48.00  754.99  8.3e-33   593   ! H91868 ys81a06.r1 Soares retina
gb_est1:AA296944    +   44.00  695.82  1.6e-29   226   ! AA296944 EST112526 Adrenal glan
gb_est2:BG967610    +   44.00  693.59  2.2e-29   360   ! BG967610 602833403F1 NIH_CGAP_C
gb_est2:BG918158    +   44.00  693.16  2.3e-29   394   ! BG918158 602821076F1 NIH_CGAP_M
gb_est2:BF786558    +   44.00  690.42  3.3e-29   697   ! BF786558 602108505F1 NIH_CGAP_K
gb_est2:BF698196    +   44.00  689.08  3.9e-29   923   ! BF698196 602125529F1 NIH_MGC_56
gb_est2:BI692031    +   44.00  688.77  4.1e-29   984   ! BI692031 603311680F1 NIH_CGAP_M
gb_est2:BE749735    -   42.00  660.76  1.5e-27   437   ! BE749735 200771 MARC 2PIG Sus s
gb_est2:BG404254    +   39.00  609.62  1.0e-24   870   ! BG404254 602420322F1 NIH_MGC_94
gb_est2:BE711179    +   37.00  583.81  2.8e-23   244   ! BE711179 RC6-HT0678-230500-011-
gb_est2:BE775078    +   37.00  582.31  3.5e-23   334   ! BE775078 IL2-UM0077-300500-089-
gb_est2:BE775099    +   37.00  582.31  3.5e-23   334   ! BE775099 IL2-UM0077-310500-0
gb_est1:AA298931    +   36.00  566.36  2.7e-22   334   ! AA298931 EST114527 Pancreas
gb_est1:BB862907    +   36.00  564.69  3.3e-22   473   ! BB862907 BB862907 RIKEN ful
gb_est2:T50828      +   34.00  535.36  1.4e-20   277   ! T50828 yb31c06.r1 Stratagen
gb_est1:AI909733    -   34.00  532.67  2.0e-20   486   ! AI909733 IL-BT223-160499-00
gb_est1:AA297274    +   31.00  488.22  6.0e-18   239   ! AA297274 EST112969 Endometr
gb_est2:BG966729    +   28.00  434.75  5.7e-15   774   ! BG966729 602834592F1 NCI_CG

seq_name: gb_est1:AL542189

seq_documentation_block:
LOCUS       AL542189               973 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL542189 LTI_FL002_PL1 Homo sapiens cDNA clone CS0DE008YG21 5 prime
            , mRNA sequence.
ACCESSION   AL542189
VERSION     AL542189.1  GI:12873988
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 973)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .973
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DE008YG21"
                     /clone_lib="LTI_FL002_PL1"
                     /lab_host="DH10B"
                     /note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
                     cDNA was primed with a NotI-oligo(dT) primer. Five prime
                     end enriched, double-stranded cDNA was digested with Not I
                     and cloned into the Not I and Eco RV sites of the
                     pCMVSPORT 6 vector.  Library was constructed by Life
                     Technologies. Contact : Feng Liang Life Technologies, a
                     division of Invitrogen 9800 Medical Center Drive Rockville
                     , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
                     fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      275 a    239 c    241 g    217 t      1 others
ORIGIN


alignment_scores:
           Quality:  193.00            Length:    193
             Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AL542189   ..

 Align seg 1/1  to: AL542189  from: 1  to: 973

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      35 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 84

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      85 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 134

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     135 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 184

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 185 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 234

  67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 235 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 284

  84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 285 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 334

 101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 335 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 384

 117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 385 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 434

 134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 435 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 484

 151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 485 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 534

 167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 535 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 584

 184 lnAlaCysArgGlyThrGluLeuAspAsp 193
     |||||||||||||||||||||||||||
 585 AGGCTTGCCGAGGGACCGAGCTTGATGAT 613

seq_name: gb_est1:AU130224

seq_documentation_block:
LOCUS       AU130224                774 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  AU130224 NT2RP3 Homo sapiens cDNA clone NT2RP3000452 5', mRNA
            sequence.
ACCESSION   AU130224
VERSION     AU130224.1  GI:10990578
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 774)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .774
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP3000452"
                     /clone_lib="NT2RP3"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                     cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT      220 a    179 c    193 g    178 t      4 others
ORIGIN


alignment_scores:
            Quality:  183.00              Length:    183
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AU130224    ..

 Align seg 1/1  to: AU130224  from: 1  to: 774

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     103 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 152

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     153 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 202

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     203 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 252

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     253 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 302

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     303 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 352

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     353 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 402

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     403 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 452

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     453 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 502

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     503 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 552

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     553 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 602

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIle 183
         |||||||||||||||||||||||||||||||||||||||||||||||||
     603 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATT 651

seq_name: gb_est2:BI821936

seq_documentation_block:
LOCUS       BI821936                926 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603039745F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5180651 5',
            mRNA sequence.
ACCESSION   BI821936
VERSION     BI821936.1  GI:15933486
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 926)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11450  row: j  column: 12
            High quality sequence stop: 882.
FEATURES             Location/Qualifiers
     source          1. .926
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5180651"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69.  Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon  cloning).  Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was  constructed by C.
                     Gruber (Invitrogen).  Research Genetics tracking  code
                     021. Note: this is a NIH_MGC Library."
BASE COUNT      264 a    218 c    242 g    202 t
ORIGIN


alignment_scores:
            Quality:  176.00            Length:    190
              Ratio:  0.931               Gaps:      1
 Percent Similarity:  99.474   Percent Identity: 99.474

alignment_block:
 US-09-613-508-2 x BI821936   ..

 Align seg 1/1  to: BI821936  from: 1  to: 926

       5 GlnGlyCysIleGluGluGlnGlyValGluAspSerAlaAsnGluAspSe 21
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
      65 CAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTC 114

      21 rValAspAlaLysProAspArgSerSerPheValProSerLeuPheSerL 38
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     115 AGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTA 164

      38 ysLysLysLysAsnValThrMetArgSerIleLysThrThrArgAspArg 54
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     165 AGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGA 214

      55 ValProThr.TyrGlnTyrAsnMetAsnPheGluLysLeuGlyLysCysI 71
         |||||||| ||||||||||||||||||||||||||||||||||||||||||
     215 GTGCCTACGATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCA 264

      71 leIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGlyValArg 87
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     265 TCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGA 314

      88 AsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheArgSerLe 104
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     315 AACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 364

     104 uGlyPheAspValIleValTyrAsnAspCysSerCysAlaLysMetGlnA 121
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     365 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAG 414

     121 spLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAlaCysPhe 137
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     415 ATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTC 464

     138 AlaCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrGlyLysAs 154
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     465 GCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGA 514

     154 pGlyValThrProIleLysAspLeuThrAlaHisPheArgGlyAspArgC 171
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     515 TGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGAT 564

     171 ysLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAlaCysArg 187
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     565 GCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGA 614

     188 GlyThrGluLeuAspAsp 193
         |||||||||||||||||
     615 GGGACCGAGCTTGATGAT 632

seq_name: gb_est1:AU126220

seq_documentation_block:
LOCUS       AU126220                809 bp    mRNA    linear   EST 23-OCT-2000
DEFINITION  AU126220 NT2RP1 Homo sapiens cDNA clone NT2RP1000352 5', mRNA
            sequence.
ACCESSION   AU126220
VERSION     AU126220.1  GI:10950936
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 809)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Kawai,Y., Ishii,S., Saito,K.,
            Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
  TITLE     HRI human cDNA project (Ota,T., Nishikawa,T., Suzuki,Y., Kawai,Y.,
            Ishii,S., Saito,K., Nakamura,Y., Nagai,T., Sugano,S., Isogai,T.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .809
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP1000352"
                     /clone_lib="NT2RP1"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pUC19FL3; mRNA from NT2 neuronal precursor
                     cells after 48-hours retinoic acid (RA) induction"
BASE COUNT      230 a    197 c    197 g    178 t      7 others
ORIGIN


alignment_scores:
            Quality:  159.00            Length:    159
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x AU126220   ..
```

```
Align seg 1/1  to: AU126220  from: 1  to: 809

      1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    119 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 168

     17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    169 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 218

     34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    219 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 268

     51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    269 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 318

     67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    319 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 368

     84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    369 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 418

    101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    419 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 468

    117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    469 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 518

    134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    519 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 568

    151 TyrGlyLysAspGlyValThrProIle 159
        |||||||||||||||||||||||||
    569 TATGGGAAAGATGGTGTCACACCAATA 595

seq_name: gb_est2:BI824214

seq_documentation_block:
LOCUS       BI824214               745 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603040584F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5181381 5',
            mRNA sequence.
ACCESSION   BI824214
VERSION     BI824214.1  GI:15935764
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 745)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11452  row: h  column: 22
            High quality sequence stop: 740.
FEATURES             Location/Qualifiers
     source          1. .745
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5181381"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69.  Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon  cloning).  Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was  constructed by C.
                     Gruber (Invitrogen).  Research Genetics tracking  code
                     021. Note: this is a NIH_MGC Library."
BASE COUNT      215 a    178 c    185 g    167 t
ORIGIN


alignment_scores:
            Quality:  147.00              Length:    147
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x BI824214    ..

 Align seg 1/1  to: BI824214  from: 1  to: 745

      1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     42 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 91

     17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     92 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 141

     34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    142 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 191

     51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    192 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 241

     67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    242 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 291

     84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    292 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 341

    101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    342 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 391

    117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    392 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 441

    134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGlu 147
        |||||||||||||||||||||||||||||||||||||||
    442 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAA 482

seq_name: gb_est2:BM150092

seq_documentation_block:
LOCUS       BM150092               532 bp    mRNA    linear   EST 30-NOV-2001
DEFINITION  TCAAP4Q11752 Pediatric acute myelogenous leukemia cell (FAB M1)
            Baylor-HGSC project=TCAA Homo sapiens cDNA clone TCAAP1175, mRNA
            sequence.
ACCESSION   BM150092
VERSION     BM150092.1  GI:17172294
```

```
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 532)
  AUTHORS   Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr.,
            Gunaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
  TITLE     Pediatric Leukemia cDNA Sequencing Project (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Dr. Judith F. Margolin
            Texas Children's Cancer Center and Human Genome Sequencing Center
            at Baylor College of Medicine
            1102 Bates, MC3-3320 Houston, TX 77030, USA
            Tel: 832-824-4536
            Fax: 832-825-4038
            Email: clones@txccc.org
            Seq primer: M13 primer.
FEATURES             Location/Qualifiers
     source          1. .532
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TCAAP1175"
                     /clone_lib="Pediatric acute myelogenous leukemia cell (FAB
                     M1) Baylor-HGSC project-TCAA"
                     /sex="male"
                     /tissue_type="leukopheresis"
                     /cell_type="myeloid cell"
                     /dev_stage="pediatric 6 years"
                     /lab_host="DH10B"
                     /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
                     First strand cDNA was primed with an anchored
                     XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGGCCGCAGGAGGAG(T)VN
                     3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
                     was primed with a BamHI-dC primer
                     [5'AGAGAGCTCGGATCCGCGGCCGCAATAATAATAAT(C) 3'].
                     Double-stranded cDNA was then digested with BamHI and XhoI
                     and directionally cloned into the BamHI and SalI sites of
                     lambda pSB vector. Library went through one round of
                     normalization. Library was constructed by Wei Yu at RIKEN
                     of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
                     Itoh M, Nagaoka S, SasakiN, Okazaki Y, Muramatsu M,
                     Schneider C, Hayashizaki Y, High efficiency selection of
                     full-length cDNA by improved biotinylated cap trapper.,
                     DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT      159 a    117 c    138 g    117 t      1 others
ORIGIN


alignment_scores:
            Quality:  144.00              Length:    144
              Ratio:  1.000                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x BM150092    ..

 Align seg 1/1  to: BM150092  from: 1  to: 532

      11 GlnGlyValGluAspSerAlaAsnGluAspSerValAspAlaLysProAs 27
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      99 CAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGA 148

      27 pArgSerSerPheValProSerLeuPheSerLysLysLysLysAsnValT 44
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     149 CCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCA 198

      44 hrMetArgSerIleLysThrThrArgAspArgValProThrTyrGlnTyr 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     199 CCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTAC 248

      61 AsnMetAsnPheGluLysLeuGlyLysCysIleIleIleAsnAsnLysAs 77
         ||||||||||||||||||||||||||||||||||||||||||||||||||

     249 AACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAA 298

      77 nPheAspLysValThrGlyMetGlyValArgAsnGlyThrAspLysAspA 94
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     299 CTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATG 348

      94 laGluAlaLeuPheLysCysPheArgSerLeuGlyPheAspValIleVal 110
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     349 CCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTC 398

     111 TyrAsnAspCysSerCysAlaLysMetGlnAspLeuLeuLysLysAlaSe 127
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     399 TATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTC 448

     127 rGluGluAspHisThrAsnAlaAlaCysPheAlaCysIleLeuLeuSerH 144
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     449 TGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTTTTAAGCC 498

     144 isGlyGluGluAsnValIleTyrGlyLysAsp 154
         |||||||||||||||||||||||||||||||
     499 ATGGAGAAGAAAATGTAATTTATGGGAAAGAT 530

seq_name: gb_est2:BM194512

seq_documentation_block:
LOCUS       BM194512                497 bp    mRNA    linear   EST 13-DEC-2001
DEFINITION  TCAAP2Q11752 Pediatric pre-B cell acute lymphoblastic leukemia
            Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCAAP1175, mRNA
            sequence.
ACCESSION   BM194512
VERSION     BM194512.1  GI:17653420
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 497)
  AUTHORS   Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr.,
            Gunaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
  TITLE     Pediatric Leukemia cDNA Sequencing Project (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Dr. Judith F. Margolin
            Texas Children's Cancer Center and Human Genome Sequencing Center
            at Baylor College of Medicine
            1102 Bates, MC3-3320 Houston, TX 77030, USA
            Tel: 832-824-4536
            Fax: 832-825-4038
            Email: clones@txccc.org
            Seq primer: M13 primer.
FEATURES             Location/Qualifiers
     source          1. .497
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TCAAP1175"
                     /clone_lib="Pediatric pre-B cell acute lymphoblastic
                     leukemia Baylor-HGSC project-TCBA"
                     /sex="male"
                     /tissue_type="leukopheresis"
                     /cell_type="pre-B cell"
                     /dev_stage="pediatric 2 years"
                     /lab_host="DH10B"
                     /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
                     First strand cDNA was primed with an anchored
                     XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGGCCGCAGGAGGAG(T)VN
                     3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
                     was primed with a BamHI-dC primer
                     [5'AGAGAGCTCGGATCCGCGGCCGCAATAATAATAAT(C) 3'].
                     Double-stranded cDNA was then digested with BamHI and XhoI
                     and directionally cloned into the BamHI and SalI sites of
                     lambda pSB vector. Library went through one round of
                     normalization. Library was constructed by Wei Yu at RIKEN
                     of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
                     Itoh M, Nagaoka S, SasakiN, Okazaki Y, Muramatsu M,
```

```
                    Schneider C, Hayashizaki Y, High efficiency selection of
                    full-length cDNA by improved biotinylated cap trapper.,
                    DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT      143 a    118 c    129 g    107 t
ORIGIN


alignment_scores:
            Quality:  142.00             Length:    142
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x BM194512   ..

 Align seg 1/1  to: BM194512  from: 1  to: 497

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      69 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 118

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     119 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 168

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     169 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 218

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     219 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 268

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     269 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 318

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     319 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 368

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     369 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 418

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     419 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 468

     134 laAlaCysPheAlaCysIleLeuLeu 142
         |||||||||||||||||||||||||
     469 CCGCCTGCTTCGCCTGCATCCTCTTA 494

seq_name: gb_est2:BG339524

seq_documentation_block:
LOCUS       BG339524               1006 bp    mRNA    linear   EST 27-FEB-2001
DEFINITION  602437368F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4555072 5',
            mRNA sequence.
ACCESSION   BG339524
VERSION     BG339524.1  GI:13145962
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1006)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLCM1255  row: h  column: 17
           High quality sequence stop: 539.
FEATURES             Location/Qualifiers
     source          1. .1006
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4555072"
                     /clone_lib="NIH_MGC_46"
                     /tissue_type="leiomyosarcoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using  ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
BASE COUNT      309 a    271 c    254 g    172 t
ORIGIN


alignment_scores:
            Quality:  141.00             Length:    141
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x BG339524   ..

 Align seg 1/1  to: BG339524  from: 1  to: 1006

       5 GlnGlyCysIleGluGluGlnGlyValGluAspSerAlaAsnGluAspSe 21
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     118 CAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTC 167

      21 rValAspAlaLysProAspArgSerSerPheValProSerLeuPheSerL 38
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     168 AGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTA 217

      38 ysLysLysLysAsnValThrMetArgSerIleLysThrThrArgAspArg 54
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     218 AGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGA 267

      55 ValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGlyLysCysIl 71
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     268 GTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCAT 317

      71 eIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGlyValArgA 88
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     318 CATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAA 367

      88 snGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheArgSerLeu 104
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     368 ACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTG 417

     105 GlyPheAspValIleValTyrAsnAspCysSerCysAlaLysMetGlnAs 121
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     418 GGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGA 467

     121 pLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAlaCysPheA 138
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     468 TCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCG 517

     138 laCysIleLeuLeuSerHisGly 145
         |||||||||||||||||||||||
```

```
    518 CCTGCATCCTCTTAAGCCATGGA 540

seq_name: gb_est2:BG025549

seq_documentation_block:
LOCUS       BG025549               830 bp    mRNA    linear   EST 24-JAN-2001
DEFINITION  602274928F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4363038 5',
            mRNA sequence.
ACCESSION   BG025549
VERSION     BG025549.1  GI:12412258
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 830)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis Staudt, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10008  row: g  column: 07
            High quality sequence stop: 669.
FEATURES             Location/Qualifiers
     source          1. .830
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4363038"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      229 a    206 c    196 g    197 t      2 others
ORIGIN


alignment_scores:
            Quality:  139.00            Length:    139
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x BG025549   ..

 Align seg 1/1  to: BG025549  from: 1  to: 830

      55 ValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGlyLysCysIl 71
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 GTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCAT 50

      71 eIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGlyValArgA 88
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 CATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAA 100

      88 snGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheArgSerLeu 104
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 ACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTG 150

     105 GlyPheAspValIleValTyrAsnAspCysSerCysAlaLysMetGlnAs 121
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 GGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGA 200

     121 pLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAlaCysPheA 138
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 TCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCG 250

     138 laCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrGlyLysAsp 154
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 CCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGAT 300

     155 GlyValThrProIleLysAspLeuThrAlaHisPheArgGlyAspArgCy 171
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 GGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATG 350

     171 sLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAlaCysArgG 188
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 CAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAG 400

     188 lyThrGluLeuAspAsp 193
         ||||||||||||||||||
     401 GGACCGAGCTTGATGAT 417

seq_name: gb_est2:BE619147

seq_documentation_block:
LOCUS       BE619147               796 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601472883F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3875733 5',
            mRNA sequence.
ACCESSION   BE619147
VERSION     BE619147.1  GI:9890085
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 796)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9635  row: f  column: 22
            High quality sequence stop: 671.
FEATURES             Location/Qualifiers
     source          1. .796
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3875733"
                     /clone_lib="NIH_MGC_68"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
BASE COUNT      225 a    169 c    223 g    179 t
ORIGIN


alignment_scores:
            Quality:  138.00            Length:    138
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x BE619147   ..

 Align seg 1/1  to: BE619147  from: 1  to: 796
```

```
     12 GlyValGluAspSerAlaAsnGluAspSerValAspAlaLysProAspAr 28
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    168 GGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCG 217

     28 gSerSerPheValProSerLeuPheSerLysLysLysLysAsnValThrM 45
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    218 GTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCA 267

     45 etArgSerIleLysThrThrArgAspArgValProThrTyrGlnTyrAsn 61
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    268 TGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAAC 317

     62 MetAsnPheGluLysLeuGlyLysCysIleIleIleAsnAsnLysAsnPh 78
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    318 ATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTT 367

     78 eAspLysValThrGlyMetGlyValArgAsnGlyThrAspLysAspAlaG 95
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    368 TGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCG 417

     95 luAlaLeuPheLysCysPheArgSerLeuGlyPheAspValIleValTyr 111
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    418 AGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTAT 467

    112 AsnAspCysSerCysAlaLysMetGlnAspLeuLeuLysLysAlaSerGl 128
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    468 AATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGA 517

    128 uGluAspHisThrAsnAlaAlaCysPheAlaCysIleLeuLeuSerHisG 145
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    518 AGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATG 567

    145 lyGluGluAsnVal 149
        |||||||||||||
    568 GAGAAGAAAATGTA 581

seq_name: gb_est2:BE891855

seq_documentation_block:
LOCUS       BE891855                961 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601435409F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920522 5',
            mRNA sequence.
ACCESSION   BE891855
VERSION     BE891855.1  GI:10351599
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 961)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9752  row: a  column: 03
            High quality sequence start: 5
            High quality sequence stop: 662.
FEATURES             Location/Qualifiers
     source          1. .961
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3920522"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      274 a    224 c    254 g    209 t
ORIGIN


alignment_scores:
            Quality:  136.00             Length:    150
              Ratio:  0.913                Gaps:      1
 Percent Similarity:  99.333   Percent Identity: 99.333

alignment_block:
 US-09-613-508-2 x BE891855   ..

 Align seg 1/1  to: BE891855  from: 1  to: 961

      1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     59 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 108

     17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    109 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 158

     34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    159 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 208

     51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    209 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 258

     67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    259 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 308

     84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    309 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 358

    101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    359 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 408

    117 aLysMetGlnAspLeu.LeuLysLysAlaSerGluGluAspHisThrAsn 133
        |||||||||||||| ||||||||||||||||||||||||||||||||||
    409 CAAGATGCAAGATCTGCTTAAAAAAAGCTTCTGAAGAGGACCATACAAAT 458

    134 AlaAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnVal 149
        |||||||||||||||||||||||||||||||||||||||||||||||||
    459 GCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTA 506

seq_name: gb_est2:BM147460

seq_documentation_block:
LOCUS       BM147460                485 bp    mRNA    linear   EST 30-NOV-2001
DEFINITION  TCAAP1Q11752 Pediatric acute myelogenous leukemia cell (FAB M1)
            Baylor-HGSC project=TCAA Homo sapiens cDNA clone TCAAP1175, mRNA
            sequence.
ACCESSION   BM147460
VERSION     BM147460.1  GI:17167317
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 485)
  AUTHORS   Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr.,
            Gunaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
  TITLE     Pediatric Leukemia cDNA Sequencing Project (2001)
```

```
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Dr. Judith F. Margolin
            Texas Children's Cancer Center and Human Genome Sequencing Center
            at Baylor College of Medicine
            1102 Bates, MC3-3320 Houston, TX 77030, USA
            Tel: 832-824-4536
            Fax: 832-825-4038
            Email: clones@txccc.org
            Seq primer: M13 primer.
FEATURES             Location/Qualifiers
     source          1. .485
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TCAAP1175"
                     /clone_lib="Pediatric acute myelogenous leukemia cell (FAB
                     M1) Baylor-HGSC project-TCAA"
                     /sex="male"
                     /tissue_type="leukopheresis"
                     /cell_type="myeloid cell"
                     /dev_stage="pediatric 6 years"
                     /lab_host="DH10B"
                     /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
                     First strand cDNA was primed with an anchored
                     XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGGCCGCAGGAGGAG(T)VN
                     3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
                     was primed with a BamH1-dC primer
                     [5'AGAGAGCTCGGATCCGCGGCCGCAATAATAATAAT(C) 3'].
                     Double-stranded cDNA was then digested with BamH1 and XhoI
                     and directionally cloned into the BamH1 and SalI sites of
                     lambda pSB vector. Library went through one round of
                     normalization. Library was constructed by Wei Yu at RIKEN
                     of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
                     Itoh M, Nagaoka S, SasakiN, Okazaki Y, Muramatsu M,
                     Schneider C, Hayashizaki Y, High efficiency selection of
                     full-length cDNA by improved biotinylated cap trapper.,
                     DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT      141 a    112 c    129 g    103 t
ORIGIN


alignment_scores:
            Quality:  134.00            Length:    134
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x BM147460   ..

 Align seg 1/1  to: BM147460  from: 1  to: 485

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      69 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 118

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     119 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 168

      34 erLeuPheSerLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||
     169 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 218

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     219 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 268

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     269 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 318

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     319 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 368

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     369 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 418

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     419 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 468

     134 la 134
         ||
     469 CC 470

seq_name: gb_est2:BF971443

seq_documentation_block:
LOCUS       BF971443                656 bp    mRNA    linear   EST 22-JAN-2001
DEFINITION  602272969F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4361176 5',
            mRNA sequence.
ACCESSION   BF971443
VERSION     BF971443.1  GI:12338658
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 656)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10003  row: i  column: 17
            High quality sequence stop: 649.
FEATURES             Location/Qualifiers
     source          1. .656
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4361176"
                     /clone_lib="NIH_MGC_84"
                     /tissue_type="adrenal cortex carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site_1:
                     NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT
                     primed. Average insert size 1.229 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      188 a    155 c    161 g    152 t
ORIGIN


alignment_scores:
            Quality:  131.00            Length:    131
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x BF971443   ..

 Align seg 1/1  to: BF971443  from: 1  to: 656

      63 AsnPheGluLysLeuGlyLysCysIleIleIleAsnAsnLysAsnPheAs 79
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       2 AATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGA 51

      79 pLysValThrGlyMetGlyValArgAsnGlyThrAspLysAspAlaGluA 96
```

```
      ||||||||||||||||||||||||||||||||||||||||||||||||||
   52 TAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGG 101

   96 laLeuPheLysCysPheArgSerLeuGlyPheAspValIleValTyrAsn 112
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  102 CGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAAT 151

  113 AspCysSerCysAlaLysMetGlnAspLeuLeuLysLysAlaSerGluGl 129
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  152 GACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGA 201

  129 uAspHisThrAsnAlaAlaCysPheAlaCysIleLeuLeuSerHisGlyG 146
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  202 GGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAG 251

  146 luGluAsnValIleTyrGlyLysAspGlyValThrProIleLysAspLeu 162
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  252 AAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTG 301

  163 ThrAlaHisPheArgGlyAspArgCysLysThrLeuLeuGluLysProLy 179
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  302 ACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAA 351

  179 sLeuPhePheIleGlnAlaCysArgGlyThrGluLeuAspAsp 193
      ||||||||||||||||||||||||||||||||||||||||||
  352 ACTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGAT 394

seq_name: gb_est2:BM459840

seq_documentation_block:
LOCUS       BM459840               1075 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6422147 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5532181
            5', mRNA sequence.
ACCESSION   BM459840
VERSION     BM459840.1  GI:18508880
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1075)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12215  row: a  column: 14
            High quality sequence stop: 545.
FEATURES             Location/Qualifiers
     source          1. .1075
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5532181"
                     /clone_lib="NIH_MGC_71"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
BASE COUNT      296 a    271 c    271 g    234 t      3 others
ORIGIN


alignment_scores:
            Quality:  123.00            Length:    123
              Ratio:  1.000               Gaps:      0
```

```
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-2 x BM459840   ..

 Align seg 1/1  to: BM459840  from: 1  to: 1075

      37 SerLysLysLysLysAsnValThrMetArgSerIleLysThrThrArgAs 53
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     190 AGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGA 239

      53 pArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGlyLysC 70
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     240 CCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAAT 289

      70 ysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGlyVal 86
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     290 GCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTT 339

      87 ArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheArgSe 103
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     340 CGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAG 389

     103 rLeuGlyPheAspValIleValTyrAsnAspCysSerCysAlaLysMetG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     390 CCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGC 439

     120 lnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAlaCys 136
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     440 AAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGC 489

     137 PheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrGlyLy 153
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     490 TTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAA 539

     153 sAspGlyValThrProIle 159
         |||||||||||||||||||
     540 AGATGGTGTCACACCAATA 558

seq_name: gb_est1:AU143139

seq_documentation_block:
LOCUS       AU143139               719 bp    mRNA    linear   EST 25-OCT-2000
DEFINITION  AU143139 Y79AA1 Homo sapiens cDNA clone Y79AA1001448 5', mRNA
            sequence.
ACCESSION   AU143139
VERSION     AU143139.1  GI:11004660
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 719)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .719
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

/clone="Y79AA1001448"
/clone_lib="Y79AA1"
/cell_type="retinoblastoma"
/cell_line="Y79"
/note="Vector: pME18SFL3"
BASE COUNT 212 a 169 c 181 g 154 t 3 others
ORIGIN

alignment_scores:
Quality: 116.00 Length: 116
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-613-508-2 x AU143139 ..

Align seg 1/1 to: AU143139 from: 1 to: 719

```
     68 GlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyMe 84
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    271 GGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 320

     84 tGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysP 101
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    321 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCT 370

    101 heArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAla 117
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    371 TCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCC 420

    118 LysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnAl 134
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    421 AAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGC 470

    134 aAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIleT 151
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    471 CGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTT 520

    151 yrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheArg 167
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    521 ATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGG 570

    168 GlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIle 183
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    571 GGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATT 618
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 00:45:44 ; Search time 2424.61 Seconds
(without alignments)
6451.747 Million cell updates/sec

Title: US-09-613-508-3
Perfect score: 1159
Sequence: 1 GCACGAGCGGATGGGTGCTA.....GGGGACCTACTCTCATGCTG 1159

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	756	65.2	827	9	AL534526	AL534526 AL534526
	2	681	58.8	913	10	BM461940	BM461940 AGENCOURT
	3	650	56.1	951	9	AU133051	AU133051 AU133051
	4	551	47.5	826	9	AU127381	AU127381 AU127381
	5	546	47.1	846	9	AU119346	AU119346 AU119346
	6	508	43.8	819	9	AU125557	AU125557 AU125557
	7	463	39.9	624	9	AV649560	AV649560 AV649560
	8	458	39.5	599	9	AW951022	AW951022 EST363092
	9	448	38.7	911	10	BE780308	BE780308 601468289
	10	446	38.5	497	9	AA984697	AA984697 am89h06.s
	11	430	37.1	624	9	AV649613	AV649613 AV649613
	12	396	34.2	751	9	AU127028	AU127028 AU127028
	13	396	34.2	946	10	BG177093	BG177093 602313329
	14	389	33.6	1007	10	BG287359	BG287359 602382123
	15	362	31.2	565	9	AL602756	AL602756 DKFZp686H
	16	356	30.7	602	10	BF966141	BF966141 602286412
	17	350	30.2	399	10	T10341	T10341 seq1254 b4H
	18	340	29.3	523	9	AL602306	AL602306 DKFZp686J
c	19	328	28.3	445	9	AA011446	AA011446 zi01g12.s
	20	304	26.2	387	9	AV648018	AV648018 AV648018
	21	270	23.3	329	9	AA298723	AA298723 EST114329
	22	237	20.4	479	9	AU127408	AU127408 AU127408
	23	236	20.4	1098	10	BG392166	BG392166 602410279
	24	223	19.2	317	9	AA297850	AA297850 EST113482
	25	213	18.4	689	12	AQ039691	AQ039691 CIT-HSP-2
	26	211	18.2	263	9	AA641754	AA641754 ns19c06.r
	27	204	17.6	965	10	BG170404	BG170404 602322734
	28	202	17.4	308	10	BM194175	BM194175 TCAAP1Q11
	29	195	16.8	414	10	BM149220	BM149220 TCAAP2Q11
	30	189	16.3	244	10	BF171567	BF171567 PCL2964 M
	31	183	15.8	411	9	AA243814	AA243814 zr67e07.r
c	32	179	15.4	424	10	BF825247	BF825247 CM1-HN001
	33	175	15.1	255	10	N84775	N84775 J0832F Huma
	34	170	14.7	352	9	AA641875	AA641875 ns19h11.r
	35	137	11.8	438	9	AA011445	AA011445 zi01g12.r
	36	127	11.0	309	9	AA298542	AA298542 EST114158
	37	123	10.6	285	10	N85243	N85243 J2957F Huma
	38	89	7.7	441	10	H29199	H29199 ym31f11.r1
c	39	80	6.9	378	9	BE178741	BE178741 PM4-HT060
	40	76	6.6	904	10	BI197761	BI197761 602756927
	41	65	5.6	115	9	AU076852	AU076852 AU076852
	42	65	5.6	704	9	AI065153	AI065153 HA0993 Hu
	43	56	4.8	1332	10	BG171387	BG171387 602322469
	44	52	4.5	435	10	BI132775	BI132775 AR035B10P
	45	49	4.2	554	10	BF104049	BF104049 601647626

ALIGNMENTS

RESULT 1
AL534526
LOCUS AL534526 827 bp mRNA linear EST 13-FEB-2001
DEFINITION AL534526 LTI_FL013_FBrn1 Homo sapiens cDNA clone CS0DF004YA06 5 prime, mRNA sequence.
ACCESSION AL534526
VERSION AL534526.1 GI:12798019
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 827)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES Location/Qualifiers
source 1. .827
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DF004YA06"
/clone_lib="LTI_FL013_FBrn1"
/dev_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"
/lab_host="DH10B"
/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville , Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 266 a 140 c 192 g 229 t

ORIGIN

Query Match 65.2%; Score 756; DB 9; Length 827;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 756; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     36 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 95

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     96 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 155

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    156 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 215

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    216 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 275

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    276 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 335

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    336 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 395

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    396 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 455

Qy    628 CTGAAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    456 CTGAAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 515

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    516 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 575

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    576 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 635

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    636 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 695

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    696 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 755

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTT 963
          |||||||||||||||||||||||||||||||||||
Db    756 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTT 791
```

RESULT 2
BM461940
LOCUS BM461940 913 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6418271 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5534002 5', mRNA sequence.
ACCESSION BM461940
VERSION BM461940.1 GI:18510980
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 913)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12219 row: m column: 11
High quality sequence stop: 671.
FEATURES Location/Qualifiers
source 1. .913
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5534002"
/clone_lib="NIH_MGC_71"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb. "
BASE COUNT 285 a 159 c 203 g 265 t 1 others
ORIGIN

Query Match 58.8%; Score 681; DB 10; Length 913;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 781; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     24 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 83

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     84 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 143

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    144 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 203

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    204 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 263

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    264 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 323

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    324 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 383

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    384 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 443

Qy    628 CTGAAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    444 CTGAAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 503

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    504 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 563

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    564 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 623
```

```
Qy     808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
           ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db     624 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTNCAAGGATGGCTCCTGGTTC 683

Qy     868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     684 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 743

Qy     928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     744 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 803

Qy     988 TTT 990
           |||
Db     804 TTT 806
```

```
RESULT   3
AU133051
LOCUS       AU133051                951 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  AU133051 NT2RP4 Homo sapiens cDNA clone NT2RP4001162 5', mRNA
            sequence.
ACCESSION   AU133051
VERSION     AU133051.1  GI:10993590
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 951)
  AUTHORS   Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J.,
            Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuho,Y. and
            Isogai,T.
  TITLE     HRI human cDNA project (Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y.,
            Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T., Sugano
            ,S., Masuho,Y., Isogai,T.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .951
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP4001162"
                     /clone_lib="NT2RP4"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                     cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT      276 a    186 c    220 g    266 t      3 others
ORIGIN
```

```
  Query Match           56.1%;  Score 650;  DB 9;  Length 951;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches  700;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 AGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCAT 65
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      19 AGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCAT 78

Qy      66 ACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGT 125
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      79 ACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGT 138

Qy     126 GAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCA 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     139 GAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCA 198

Qy     186 CCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGT 245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     199 CCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGT 258

Qy     246 GGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGA 305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     259 GGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGA 318

Qy     306 CTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTAT 365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     319 CTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTAT 378

Qy     366 AATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGA 425
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     379 AATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTGCAGA 438

Qy     426 TGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAA 485
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     439 TGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAA 498

Qy     486 AAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCA 545
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     499 AAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCA 558

Qy     546 CAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTT 605
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     559 CAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTT 618

Qy     606 TGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTG 665
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     619 TGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTG 678

Qy     666 TAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGG 706
           |||||||||||||||||||||||||||||||||||||||||
Db     679 TAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGG 719
```

```
RESULT   4
AU127381
LOCUS       AU127381                826 bp    mRNA    linear   EST 23-OCT-2000
DEFINITION  AU127381 NT2RP2 Homo sapiens cDNA clone NT2RP2001169 5', mRNA
            sequence.
ACCESSION   AU127381
VERSION     AU127381.1  GI:10952097
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 826)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .826
```

```
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP2001169"
                     /clone_lib="NT2RP2"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                     cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT      254 a    152 c    188 g    226 t      6 others
ORIGIN


  Query Match             47.5%;  Score 551;  DB 9;  Length 826;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-279;
  Matches  551;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 AGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCAT 65
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16 AGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCAT 75

Qy      66 ACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGT 125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      76 ACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGT 135

Qy     126 GAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCA 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     136 GAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCA 195

Qy     186 CCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGT 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     196 CCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGT 255

Qy     246 GGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGA 305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     256 GGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGA 315

Qy     306 CTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTAT 365
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     316 CTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTAT 375

Qy     366 AATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGA 425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     376 AATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGA 435

Qy     426 TGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAA 485
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     436 TGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAA 495

Qy     486 AAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCA 545
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     496 AAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCA 555

Qy     546 CAGCAAAAGGA 556
          ||||||||||||
Db     556 CAGCAAAAGGA 566


RESULT   5
AU119346
LOCUS       AU119346                846 bp    mRNA    linear   EST 19-OCT-2000
DEFINITION  AU119346 HEMBA1 Homo sapiens cDNA clone HEMBA1005587 5', mRNA
            sequence.
ACCESSION   AU119346
VERSION     AU119346.1  GI:10934581
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 846)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .846
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEMBA1005587"
                     /clone_lib="HEMBA1"
                     /tissue_type="whole embryo, mainly head"
                     /dev_stage="embryo, 10 weeks"
                     /note="Vector: pME18SFL3"
BASE COUNT      266 a    150 c    200 g    226 t      4 others
ORIGIN


  Query Match             47.1%;  Score 546;  DB 9;  Length 846;
  Best Local Similarity   99.8%;  Pred. No. 6.1e-277;
  Matches  596;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      53 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 112

Qy     268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     113 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 172

Qy     328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     173 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 232

Qy     388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     233 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 292

Qy     448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     293 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAGTGATCTTACACGTGAAGAA 352

Qy     508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     353 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 412

Qy     568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     413 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 472

Qy     628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     473 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 532

Qy     688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     533 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 592

Qy     748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTG 804
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     593 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTG 649


RESULT   6
AU125557
```

```
LOCUS       AU125557                819 bp    mRNA    linear   EST 23-OCT-2000
DEFINITION  AU125557 NT2RM4 Homo sapiens cDNA clone NT2RM4001791 5', mRNA
            sequence.
ACCESSION   AU125557
VERSION     AU125557.1  GI:10950273
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 819)
  AUTHORS   Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,
            Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,
            Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki
            ,Y., Sugano,S., Isogai,T.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .819
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RM4001791"
                     /clone_lib="NT2RM4"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
                     precursor cells"
BASE COUNT      244 a    149 c    168 g    254 t      4 others
ORIGIN
```

```
  Query Match           43.8%;  Score 508;  DB 9;  Length 819;
  Best Local Similarity  100.0%;  Pred. No. 7.2e-257;
  Matches  508;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     611 CAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAA 670
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAA 60

Qy     671 GTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACT 730
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACT 120

Qy     731 GTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGG 790
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGG 180

Qy     791 AGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAA 850
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAA 240

Qy     851 AGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGC 910
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGC 300

Qy     911 TTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCT 970
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCT 360

Qy     971 TTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCA 1030
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCA 420

Qy    1031 CAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGT 1090
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGT 480

Qy    1091 ATGCCAAGTGAGAAGATGGTATATTTGG 1118
           ||||||||||||||||||||||||||||
Db     481 ATGCCAAGTGAGAAGATGGTATATTTGG 508
```

```
RESULT   7
AV649560
LOCUS       AV649560                624 bp    mRNA    linear   EST 15-JAN-2002
DEFINITION  AV649560 GLC Homo sapiens cDNA clone GLCBVC04 3', mRNA sequence.
ACCESSION   AV649560
VERSION     AV649560.1  GI:9870574
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
            Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
            Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
            Hu,G., Gu,J., Chen,Z. and Han,Z.
  TITLE     Insight into hepatocellular carcinogenesis at transcriptome level
            by comparing gene expression profiles of hepatocellular carcinoma
            with those of corresponding noncancerous liver
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
  MEDLINE   21625106
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="GLCBVC04"
                     /clone_lib="GLC"
                     /tissue_type="corresponding non cancerous liver tissue"
                     /dev_stage="Adult"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      182 a    116 c    156 g    166 t      4 others
ORIGIN
```

```
  Query Match           39.9%;  Score 463;  DB 9;  Length 624;
  Best Local Similarity  99.8%;  Pred. No. 4.3e-233;
  Matches  513;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      37 GAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTG 96
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      30 GAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTG 89

Qy      97 GTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGC 156
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      90 GTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGC 149

Qy     157 GCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAG 216
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     150 GCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCTCTTCCCCCATTCTCATTAATAAAG 209

Qy     217 GTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCA 276
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    210 GTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCA 269

Qy    277 AAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAA 336
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    270 AAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAA 329

Qy    337 ATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGC 396
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    330 ATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGC 389

Qy    397 ACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTC 456
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    390 ACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTC 449

Qy    457 AGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAA 516
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    450 AGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAA 509

Qy    517 TTGATGCGTGATGTTTCTAAAGAAGATCACAGCA 550
          |||||||||||||||||||||||||||||||||
Db    510 TTGATGCGTGATGTTTCTAAAGAAGATCACAGCA 543


RESULT   8
AW951022
LOCUS       AW951022                599 bp    mRNA    linear   EST 01-JUN-2000
DEFINITION  EST363092 MAGE resequences, MAGA Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW951022
VERSION     AW951022.1  GI:8140687
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 599)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
            ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 22
            Seq primer: Reverse.
FEATURES             Location/Qualifiers
     source          1. .599
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGA"
                     /note="Vector: pBluescriptSKm"
BASE COUNT      218 a     85 c    135 g    160 t      1 others
ORIGIN


  Query Match           39.5%;  Score 458;  DB 9;  Length 599;
  Best Local Similarity  100.0%;  Pred. No. 1.9e-230;
  Matches  458;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 82

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     83 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 142

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    143 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 202

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 262

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 322

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    323 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 382

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    383 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 442

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTG 665
          |||||||||||||||||||||||||||||||||||||
Db    443 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTG 480


RESULT   9
BE780308
LOCUS       BE780308                911 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601468289F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3871390 5',
            mRNA sequence.
ACCESSION   BE780308
VERSION     BE780308.1  GI:10201506
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 911)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9624  row: a  column: 23
            High quality sequence stop: 768.
FEATURES             Location/Qualifiers
     source          1. .911
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3871390"
                     /clone_lib="NIH_MGC_67"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      287 a    175 c    205 g    244 t
ORIGIN


  Query Match           38.7%;  Score 448;  DB 10;  Length 911;
  Best Local Similarity  99.8%;  Pred. No. 3.6e-225;
  Matches  568;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy    235 GAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGC 294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     14 GAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGC 73
```

```
Qy    295 GAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATG 354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74 GAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATG 133

Qy    355 GGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGG 414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    134 GGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGG 193

Qy    415 TCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAA 474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    194 TCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAA 253

Qy    475 GTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCT 534
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    254 GTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCT 313

Qy    535 AAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAA 594
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    314 AAAGAAGATCACAGCAAAAGGAGCAG-TTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAA 372

Qy    595 GGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGA 654
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    373 GGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGA 432

Qy    655 GGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGT 714
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    433 GGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGT 492

Qy    715 GGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGT 774
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    493 GGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGT 552

Qy    775 CATAAAATACCAGTGGAGGCCGACTTCTT 803
          |||||||||||||||||||||||||||||
Db    553 CATAAAATACCAGTGGAGGCCGACTTCTT 581
```

```
RESULT  10
AA984697
LOCUS       AA984697                497 bp    mRNA    linear   EST 27-MAY-1998
DEFINITION  am89h06.s1 Stratagene schizo brain S11 Homo sapiens cDNA clone
            IMAGE:1630331 3' similar to SW:ICE3_HUMAN P42574 APOPAIN PRECURSOR
            ;, mRNA sequence.
ACCESSION   AA984697
VERSION     AA984697.1  GI:3163222
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 497)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 372.
FEATURES             Location/Qualifiers
     source          1. .497
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1630331"
                     /clone_lib="Stratagene schizo brain S11"
                     /sex="male"
                     /tissue_type="schizophrenic brain S-11 frontal lobe"
                     /dev_stage="34 years old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Vector: Bluescript SK-; Site_1: EcoRI;  Library
                     constructed from S-11 frontal lobe, male, 34 years old,
                     50% caucasian, 50% Aleutian.  Schizophrenic suicide.
                     Random primed into EcoRI site of ZAP II Vector.  Mass
                     excised. Custom library. Avg insert length 1.4kb.
                     Material obtained by Johnston N., Torrey, E.F., Yolken R.,
                     and the Stanley Neuropathology Consortium - Analysis of
                     RNAs from the Brains of Individuals with Psychiatric
                     Diseases (Unpublished) Stanley Neurovirology Laboratory,
                     Johns Hopkins School of Medicine, Baltimore MD."
BASE COUNT      153 a    101 c    113 g    130 t
ORIGIN
```

```
  Query Match           38.5%;  Score 446;  DB 9;  Length 497;
  Best Local Similarity   99.8%;  Pred. No. 4.2e-224;
  Matches  496;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     50 AGTGCTCGCAGCTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCT 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AGTGCTCGCAGCTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCT 60

Qy    110 TGAAATCCCAGGCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCG 169
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGAAATCCCAGGCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCG 120

Qy    170 GTGGGTGTGCCCTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGA 229
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    121 GTGGGTGTGCCCTGCACCTGCCTCTTCCCCCATTCTCATTAATAAAGGTATCCATGGAGA 180

Qy    230 ACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATG 289
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATG 240

Qy    290 GAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTG 349
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTG 300

Qy    350 AGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACAT 409
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACAT 360

Qy    410 CTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAAT 469
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAAT 420

Qy    470 ATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATG 529
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATG 480

Qy    530 TTTCTAAAGAAGATCAC 546
          |||||||||||||||||
Db    481 TTTCTAAAGAAGATCAC 497
```

```
RESULT  11
AV649613
LOCUS       AV649613                624 bp    mRNA    linear   EST 15-JAN-2002
DEFINITION  AV649613 GLC Homo sapiens cDNA clone GLCBVG11 3', mRNA sequence.
ACCESSION   AV649613
VERSION     AV649613.1  GI:9870627
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 624)
```

```
  AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
            Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
            Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
            Hu,G., Gu,J., Chen,Z. and Han,Z.
  TITLE     Insight into hepatocellular carcinogenesis at transcriptome level
            by comparing gene expression profiles of hepatocellular carcinoma
            with those of corresponding noncancerous liver
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
  MEDLINE   21625106
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="GLCBVG11"
                     /clone_lib="GLC"
                     /tissue_type="corresponding non cancerous liver tissue"
                     /dev_stage="Adult"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      181 a    116 c    157 g    163 t      7 others
ORIGIN
```

```
  Query Match             37.1%;  Score 430;  DB 9;  Length 624;
  Best Local Similarity   99.8%;  Pred. No. 1.2e-215;
  Matches  480;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      37 GAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTG 96
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      30 GAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTG 89

Qy      97 GTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGC 156
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      90 GTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGC 149

Qy     157 GCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAG 216
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     150 GCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCTCTTCCCCCATTCTCATTAATAAAG 209

Qy     217 GTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCA 276
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     210 GTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCA 269

Qy     277 AAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAA 336
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     270 AAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAA 329

Qy     337 ATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGC 396
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     330 ATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGC 389

Qy     397 ACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTC 456
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     390 ACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTC 449

Qy     457 AGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAA 516
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     450 AGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAA 509

Qy     517 T 517
           |
Db     510 T 510
```

```
RESULT  12
AU127028
LOCUS       AU127028                751 bp    mRNA    linear   EST 23-OCT-2000
DEFINITION  AU127028 NT2RP2 Homo sapiens cDNA clone NT2RP2000417 5', mRNA
            sequence.
ACCESSION   AU127028
VERSION     AU127028.1  GI:10951744
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 751)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .751
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP2000417"
                     /clone_lib="NT2RP2"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                     cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT      242 a    129 c    175 g    194 t     11 others
ORIGIN
```

```
  Query Match             34.2%;  Score 396;  DB 9;  Length 751;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-197;
  Matches  446;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 141

Qy     268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 201

Qy     328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     202 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 261

Qy     388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     262 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 321

Qy     448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     322 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 381

Qy     508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     382 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 441

Qy     568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    442 TGTGTGCTTCTGANCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 501

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGA 654
          ||||||||||||||||||||||||||||
Db    502 CTGAAAAAAATAACAAACTTTTTCAGA 528


RESULT  13
BG177093
LOCUS       BG177093                 946 bp    mRNA    linear   EST 06-FEB-2001
DEFINITION  602313329F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4419175 5',
            mRNA sequence.
ACCESSION   BG177093
VERSION     BG177093.1  GI:12683796
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 946)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis Staudt, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10154  row: j  column: 08
            High quality sequence stop: 644.
FEATURES             Location/Qualifiers
     source          1. .946
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4419175"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      271 a    237 c    226 g    212 t
ORIGIN


  Query Match           34.2%;  Score 396;  DB 10;  Length 946;
  Best Local Similarity   99.8%;  Pred. No. 1e-197;
  Matches  446;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     37 GAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTG 96
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     42 GAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTG 101

Qy     97 GTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGC 156
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    102 GTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGC 161

Qy    157 GCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAG 216
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    162 GCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCTCTTCCCCCATTCTCATTAATAAAG 221

Qy    217 GTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCA 276
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 GTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCA 281

Qy    277 AAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAA 336
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    282 AAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAA 341

Qy    337 ATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGC 396
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    342 ATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGC 401

Qy    397 ACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTC 456
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    402 ACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTC 461

Qy    457 AGAAACTTGAAATATGAAGTCAGGAAT 483
          |||||||||||||||||||||||||||
Db    462 AGAAACTTGAAATATGAAGTCAGGAAT 488


RESULT  14
BG287359
LOCUS       BG287359                 1007 bp    mRNA    linear   EST 21-FEB-2001
DEFINITION  602382123F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4499885 5',
            mRNA sequence.
ACCESSION   BG287359
VERSION     BG287359.1  GI:13041121
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1007)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10364  row: m  column: 06
            High quality sequence stop: 655.
FEATURES             Location/Qualifiers
     source          1. .1007
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4499885"
                     /clone_lib="NIH_MGC_93"
                     /tissue_type="transitional cell papilloma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.7 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      265 a    213 c    235 g    294 t
ORIGIN


  Query Match           33.6%;  Score 389;  DB 10;  Length 1007;
  Best Local Similarity   100.0%;  Pred. No. 5.2e-194;
  Matches  389;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    694 TTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGT 753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGT 60

Qy    754 GTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATAC 813
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATAC 120

Qy    814 TCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAG 873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    121 TCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAG 180

Qy    874 TCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACC 933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACC 240

Qy    934 CGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCAT 993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCAT 300

Qy    994 GCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1053
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 360

Qy   1054 TAAAGAAATGGTTGGTTGGTGGTTTTTTT 1082
          |||||||||||||||||||||||||||||
Db    361 TAAAGAAATGGTTGGTTGGTGGTTTTTTT 389


RESULT  15
AL602756
LOCUS       AL602756                565 bp    mRNA    linear   EST 14-AUG-2001
DEFINITION  DKFZp686H1516_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686H1516 5', mRNA sequence.
ACCESSION   AL602756
VERSION     AL602756.1  GI:15166262
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 565)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,W., Weil,B. and Wiemann,S.
  TITLE     EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B. and Wiemann,S.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Wambutt R
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686H1516) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .565
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686H1516"
                     /clone_lib="686 (synonym: hlcc3)"
                     /tissue_type="human skeletal muscle"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
BASE COUNT      174 a    102 c    136 g    153 t
ORIGIN


  Query Match           31.2%;  Score 362;  DB 9;  Length 565;
  Best Local Similarity  100.0%;  Pred. No. 9.8e-180;
  Matches  362;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    204 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 263

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    264 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 323

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    324 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 383

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    384 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 443

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    444 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 503

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    504 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 563

Qy    568 TG 569
          ||
Db    564 TG 565


Search completed: August 16, 2002, 00:45:52
Job time: 6208 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 01:43:01 ; Search time 75.09 Seconds
(without alignments)
3791.307 Million cell updates/sec

Title: US-09-613-508-3
Perfect score: 1159
Sequence: 1 GCACGAGCGGATGGGTGCTA.....GGGGACCTACTCTCATGCTG 1159

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1159	100.0	1159	3	US-08-462-969B-3	Sequence 3, Appli
	2	1118	96.5	1158	3	US-08-591-605-1	Sequence 1, Appli
	3	1004	86.6	2635	4	US-09-484-617-3	Sequence 3, Appli
	4	860	74.2	1001	3	US-08-964-308-5	Sequence 5, Appli
	5	860	74.2	1001	3	US-08-964-313-5	Sequence 5, Appli
	6	860	74.2	1001	4	US-09-069-138-5	Sequence 5, Appli
	7	809	69.8	1001	3	US-08-964-308-9	Sequence 9, Appli
	8	809	69.8	1001	3	US-08-964-313-9	Sequence 9, Appli
	9	809	69.8	1001	4	US-09-069-138-9	Sequence 9, Appli
	10	32	2.8	53	2	US-08-800-007A-1	Sequence 1, Appli
	11	32	2.8	53	2	US-08-800-007A-5	Sequence 5, Appli
c	12	31	2.7	51	2	US-08-800-007A-4	Sequence 4, Appli
c	13	31	2.7	51	2	US-08-800-007A-8	Sequence 8, Appli
	14	30	2.6	90	4	US-09-208-966-20	Sequence 20, Appl
	15	30	2.6	90	4	US-09-208-966-32	Sequence 32, Appl
c	16	29	2.5	52	2	US-08-800-007A-2	Sequence 2, Appli
c	17	29	2.5	52	2	US-08-800-007A-6	Sequence 6, Appli
	18	29	2.5	1177	2	US-08-890-542A-1	Sequence 1, Appli
	19	29	2.5	1177	2	US-08-890-542A-3	Sequence 3, Appli
	20	29	2.5	1297	4	US-09-484-617-10	Sequence 10, Appl
	21	28	2.4	47	2	US-08-800-007A-3	Sequence 3, Appli
	22	28	2.4	47	2	US-08-800-007A-7	Sequence 7, Appli
	23	27	2.3	27	4	US-09-484-617-6	Sequence 6, Appli
	24	25	2.2	31	3	US-08-462-969B-7	Sequence 7, Appli
	25	25	2.2	85	4	US-09-208-966-24	Sequence 24, Appl
	26	25	2.2	85	4	US-09-208-966-36	Sequence 36, Appl
c	27	23	2.0	53	3	US-08-462-969B-12	Sequence 12, Appl
c	28	22	1.9	74	4	US-09-208-966-33	Sequence 33, Appl
c	29	22	1.9	75	4	US-09-208-966-22	Sequence 22, Appl
	30	21	1.8	21	4	US-09-484-617-4	Sequence 4, Appli
c	31	21	1.8	31	3	US-08-462-969B-8	Sequence 8, Appli
c	32	21	1.8	37	4	US-09-208-966-23	Sequence 23, Appl
c	33	21	1.8	37	4	US-09-208-966-35	Sequence 35, Appl
	34	21	1.8	72	4	US-09-208-966-21	Sequence 21, Appl
	35	21	1.8	72	4	US-09-208-966-34	Sequence 34, Appl
c	36	20	1.7	20	4	US-09-484-617-18	Sequence 18, Appl
c	37	20	1.7	20	4	US-09-484-617-19	Sequence 19, Appl
c	38	20	1.7	20	4	US-09-484-617-20	Sequence 20, Appl
c	39	20	1.7	20	4	US-09-484-617-21	Sequence 21, Appl
c	40	20	1.7	20	4	US-09-484-617-22	Sequence 22, Appl
c	41	20	1.7	20	4	US-09-484-617-23	Sequence 23, Appl
c	42	20	1.7	20	4	US-09-484-617-24	Sequence 24, Appl
c	43	20	1.7	20	4	US-09-484-617-25	Sequence 25, Appl
c	44	20	1.7	20	4	US-09-484-617-26	Sequence 26, Appl
c	45	20	1.7	20	4	US-09-484-617-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-08-462-969B-3
; Sequence 3, Application US/08462969B
; Patent No. 6087150
; GENERAL INFORMATION:
; APPLICANT: He, Wei-Wu et al.
; TITLE OF INVENTION: Interleukin-1 Beta Converting Enzyme
; TITLE OF INVENTION: Like Apoptosis Protease 3 and 4
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Ave.
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,969B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/334,251
; FILING DATE: 11-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF140P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1159 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-462-969B-3

Query Match 100.0%; Score 1159; DB 3; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60

```
Db      1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60

Qy     61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120
Db     61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120

Qy    121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180
Db    121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180

Qy    181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240
Db    181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240

Qy    241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300
Db    241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300

Qy    301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360
Db    301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360

Qy    361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420
Db    361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420

Qy    421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480
Db    421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480

Qy    481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540
Db    481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540

Qy    541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600
Db    541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600

Qy    601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660
Db    601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660

Qy    661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720
Db    661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720

Qy    721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780
Db    721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780

Qy    781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840
Db    781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840

Qy    841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900
Db    841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900

Qy    901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960
Db    901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960

Qy    961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020
Db    961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020

Qy   1021 TCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080
Db   1021 TCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080

Qy   1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTG 1140

Db   1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTG 1140

Qy   1141 GGGACCTACTCTCATGCTG 1159
Db   1141 GGGACCTACTCTCATGCTG 1159
```

RESULT 2
US-08-591-605-1
; Sequence 1, Application US/08591605
; Patent No. 6060238
; GENERAL INFORMATION:
; APPLICANT: Dixit, Vishva M.
; TITLE OF INVENTION: METHOD AND COMPOSITION FOR REGULATING
; TITLE OF INVENTION: APOPTOSIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,605
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: KONSKI, ANTOINETTE F.
; REGISTRATION NUMBER: 34,202
; REFERENCE/DOCKET NUMBER: 20344-21036.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141 MRSNFOERS SFO
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1158 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 224..1054
; OTHER INFORMATION: /product= "Yama peptide sequence"
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 704..718
; OTHER INFORMATION: /note= "region coding for
; OTHER INFORMATION: pentapeptide, wherein cysteine is catalytic"
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(710..712, "atg")
; OTHER INFORMATION: /note= "Mutation of catalytic cysteine to methionine"
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(710..712, "gcc")
; OTHER INFORMATION: /note= "mutation of catalytic cysteine to alanine"
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(710..712, "gcg")
; OTHER INFORMATION: /note= "mutation of catalytic cysteine to alanine"
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(710..712, "gct")
; OTHER INFORMATION: /note= "mutation of catalytic cysteine to alanine"
; FEATURE:
; NAME/KEY: mutation

```
;      LOCATION:  replace(710..712, "gca")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
US-08-591-605-1


  Query Match             96.5%;  Score 1118;  DB 3;  Length 1158;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1118;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 61

Qy      61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 121

Qy     121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 181

Qy     181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 241

Qy     241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 301

Qy     301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 361

Qy     361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 421

Qy     421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     422 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 481

Qy     481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     482 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 541

Qy     541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     542 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 601

Qy     601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     602 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 661

Qy     661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     662 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 721

Qy     721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     722 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 781

Qy     781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     782 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 841

Qy     841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     842 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 901

Qy     901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     902 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 961
```

```
Qy     961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     962 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1021

Qy    1021 TCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1022 TCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1081

Qy    1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGG 1118
           |||||||||||||||||||||||||||||||||||||
Db    1082 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGG 1119


RESULT   3
US-09-484-617-3
; Sequence 3, Application US/09484617
; Patent No. 6303374
; GENERAL INFORMATION:
;  APPLICANT: Hong Zhang
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF CASPASE 3 EXPRESSION
;  FILE REFERENCE: RTS-0103
;  CURRENT APPLICATION NUMBER: US/09/484,617
;  CURRENT FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 176
; SEQ ID NO 3
;   LENGTH: 2635
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (225)...(1058)
US-09-484-617-3


  Query Match             86.6%;  Score 1004;  DB 4;  Length 2635;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1104;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      13 GGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTG 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 gggtgctattgtgaggcggttgtagaagagtttcgtgagtgctcgcagctcatacctgtg 74

Qy      73 GCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGT 132
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 gctgtgtatccgtggccacagctggttggcgtcgccttgaaatcccaggccgtgaggagt 134

Qy     133 TAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 tagcgagccctgctcacactcggcgctctggttttcggtgggtgtgccctgcacctgcct 194

Qy     193 CTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCA 252
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     195 cttccccattctcattaataaaggtatccatggagaacactgaaaactcagtggattca 254

Qy     253 AAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGA 312
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     255 aaatccattaaaaatttggaaccaaagatcatacatggaagcgaatcaatggactctgga 314

Qy     313 ATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     315 atatccctggacaacagttataaaatggattatcctgagatgggtttatgtataataatt 374

Qy     373 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGAT 432
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     375 aataataagaattttcataaaagcactggaatgacatctcggtctggtacagatgtcgat 434

Qy     433 GCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGAT 492
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     435 gcagcaaacctcagggaaacattcagaaacttgaaatatgaagtcaggaataaaaatgat 494

Qy     493 CTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAA 552
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    495 cttacacgtgaagaaattgtggaattgatgcgtgatgtttctaaagaagatcacagcaaa 554

Qy    553 AGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACA 612
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    555 aggagcagttttgtttgtgtgcttctgagccatggtgaagaaggaataatttttggaaca 614

Qy    613 AATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 672
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    615 aatggacctgttgacctgaaaaaaataacaaacttttcagaggggatcgttgtagaagt 674

Qy    673 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGT 732
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    675 ctaactggaaaacccaaacttttcattattcaggcctgccgtggtacagaactggactgt 734

Qy    733 GGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAG 792
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    735 ggcattgagacagacagtggtgttgatgatgacatggcgtgtcataaaatataccagtggat 794

Qy    793 GCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAG 852
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    795 gccgacttcttgtatgcatactccacagcacctggttattattcttggcgaaattcaaag 854

Qy    853 GATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTT 912
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    855 gatggctcctggttcatccagtcgctttgtgccatgctgaaacagtatgccgacaagctt 914

Qy    913 GAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 972
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    915 gaatttatgcacattcttacccgggttaaccgaaaggtggcaacagaatttgagtccttt 974

Qy    973 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACA 1032
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    975 tcctttgacgctacttttcatgcaaagaaacagattccatgtattgtttccatgctcaca 1034

Qy   1033 AAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTAT 1092
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1035 aaagaactctatttttatcactaaagaaatggttggttggtggttttttttagtttgtat 1094

Qy   1093 GCCAAGTGAGAAGATGGTATATTTGG 1118
          |||||||||||||||||||||||||
Db   1095 gccaagtgagaagatggtatatttgg 1120


RESULT   4
US-08-964-308-5
; Sequence 5, Application US/08964308
; Patent No. 6066715
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBERT J. NORTH - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,308
;      FILING DATE:  04-NOV-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  NORTH, ROBERT J
;      REGISTRATION NUMBER:  27,366
;      REFERENCE/DOCKET NUMBER:  19840 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-7262
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-308-5


  Query Match             74.2%;  Score 860;  DB 3;  Length 1001;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  910;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    529 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy    988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy   1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy   1108 GGTATATTTGG 1118
          ||||||||||||
Db    949 GGTATATTTGG 959


RESULT   5
US-08-964-313-5
; Sequence 5, Application US/08964313
; Patent No. 6114132
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  GRESSER, MICHAEL
;    APPLICANT:  KENNEDY, BRIAN
;    APPLICANT:  NICHOLSON, DONALD
;    APPLICANT:  RAMACHANDRAN, CHIDAMBARAN
;    APPLICANT:  SKOREY, KATHRYN
;    APPLICANT:  FORD-HUTCHINSON, ANTHONY
;    TITLE OF INVENTION:  PHOSPHATASE BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,313
;      FILING DATE:  04-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/030,408
;      FILING DATE:  04-NOV-1996
;      APPLICATION NUMBER:  PCT/CA97/00825
;      FILING DATE:  03 -NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19824Y
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-313-5


  Query Match             74.2%;  Score 860;  DB 3;  Length 1001;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  910;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    529 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy    988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy   1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy   1108 GGTATATTTGG 1118
          ||||||||||||
Db    949 GGTATATTTGG 959


RESULT   6
US-09-069-138-5
; Sequence 5, Application US/09069138
; Patent No. 6348572
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  DUFRESNE, CLAUDE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  LEBLANC, YVES
;    APPLICANT:  ROY, PATRICK
;    APPLICANT:  YOUNG, ROBERT N.
;    APPLICANT:  ZAMBONI, ROBERT
```

```
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/069,138
;      FILING DATE:  29-APR-1998
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19840YIA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-069-138-5


  Query Match             74.2%;  Score 860;  DB 4;  Length 1001;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  910;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy     268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy     328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy     388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy     448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy     508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy     568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy     628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy     688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     529 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy     748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy     808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy     868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy     928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy     988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy    1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy    1108 GGTATATTTGG 1118
           ||||||||||||
Db     949 GGTATATTTGG 959


RESULT   7
US-08-964-308-9
; Sequence 9, Application US/08964308
; Patent No. 6066715
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBERT J. NORTH - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,308
;      FILING DATE:  04-NOV-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  NORTH, ROBERT J
;      REGISTRATION NUMBER:  27,366
;      REFERENCE/DOCKET NUMBER:  19840 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-7262
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-308-9

  Query Match             69.8%;  Score 809;  DB 3;  Length 1001;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches  909;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    529 AAACTTTTCATTATTCAGGCCTCCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy    988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy   1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy   1108 GGTATATTTGG 1118
          ||||||||||||
Db    949 GGTATATTTGG 959
```

```
RESULT   8
US-08-964-313-9
; Sequence 9, Application US/08964313
; Patent No. 6114132
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  GRESSER, MICHAEL
;    APPLICANT:  KENNEDY, BRIAN
;    APPLICANT:  NICHOLSON, DONALD
;    APPLICANT:  RAMACHANDRAN, CHIDAMBARAN
;    APPLICANT:  SKOREY, KATHRYN
;    APPLICANT:  FORD-HUTCHINSON, ANTHONY
;    TITLE OF INVENTION:  PHOSPHATASE BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,313
;      FILING DATE:  04-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/030,408
;      FILING DATE:  04-NOV-1996
;      APPLICATION NUMBER:  PCT/CA97/00825
;      FILING DATE:  03 -NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19824Y
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-313-9

  Query Match             69.8%;  Score 809;  DB 3;  Length 1001;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches  909;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    529 AAACTTTTCATTATTCAGGCCTCCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy    988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy   1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy   1108 GGTATATTTGG 1118
          ||||||||||| 
Db    949 GGTATATTTGG 959
```

```
RESULT   9
US-09-069-138-9
; Sequence 9, Application US/09069138
; Patent No. 6348572
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  DUFRESNE, CLAUDE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  LEBLANC, YVES
;    APPLICANT:  ROY, PATRICK
;    APPLICANT:  YOUNG, ROBERT N.
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/069,138
;      FILING DATE:  29-APR-1998
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19840YIA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-069-138-9
```

```
  Query Match             69.8%;  Score 809;  DB 4;  Length 1001;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches  909;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    529 AAACTTTTCATTATTCAGGCCTCCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy    988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy   1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy   1108 GGTATATTTGG 1118
          ||||||||||||
Db    949 GGTATATTTGG 959
```

```
RESULT  10
US-08-800-007A-1
; Sequence 1, Application US/08800007A
; Patent No. 5834228
;  GENERAL INFORMATION:
;    APPLICANT:  Becker, Joseph
;    APPLICANT:  Nicholson, Donald
;    APPLICANT:  Rotonda, Jennifer
;    APPLICANT:  Thornberry, Nancy
;    APPLICANT:  Fazil, Kimberly
;    APPLICANT:  Gallant, Michel
;    APPLICANT:  Gareau, Yves
;    APPLICANT:  Labelle, Marc
;    APPLICANT:  Peterson, Erin
;    APPLICANT:  Rasper, Dita
;    TITLE OF INVENTION:  CRYSTAL STRUCTURE OF APOPAIN
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  JOSEPH A. COPPOLA - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/800,007A
;      FILING DATE:  13-FEB-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COPPOLA, JOSEPH A
;      REGISTRATION NUMBER:  38,413
;      REFERENCE/DOCKET NUMBER:  19644
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-6734
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  53 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-800-007A-1


  Query Match             2.8%;  Score 32;  DB 2;  Length 53;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-06;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    309 TGGAATATCCCTGGACAACAGTTATAAAATGG 340
          ||||||||||||||||||||||||||||||||
Db     22 TGGAATATCCCTGGACAACAGTTATAAAATGG 53


RESULT  11
US-08-800-007A-5
; Sequence 5, Application US/08800007A
; Patent No. 5834228
;  GENERAL INFORMATION:
;    APPLICANT:  Becker, Joseph
;    APPLICANT:  Nicholson, Donald
;    APPLICANT:  Rotonda, Jennifer
;    APPLICANT:  Thornberry, Nancy
;    APPLICANT:  Fazil, Kimberly
;    APPLICANT:  Gallant, Michel
;    APPLICANT:  Gareau, Yves
;    APPLICANT:  Labelle, Marc
;    APPLICANT:  Peterson, Erin
;    APPLICANT:  Rasper, Dita
;    TITLE OF INVENTION:  CRYSTAL STRUCTURE OF APOPAIN
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  JOSEPH A. COPPOLA - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/800,007A
;      FILING DATE:  13-FEB-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COPPOLA, JOSEPH A
;      REGISTRATION NUMBER:  38,413
;      REFERENCE/DOCKET NUMBER:  19644
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-6734
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  53 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-800-007A-5


  Query Match             2.8%;  Score 32;  DB 2;  Length 53;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-06;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy     309 TGGAATATCCCTGGACAACAGTTATAAAATGG 340
           ||||||||||||||||||||||||||||||||
Db      22 TGGAATATCCCTGGACAACAGTTATAAAATGG 53


RESULT  12
US-08-800-007A-4/c
; Sequence 4, Application US/08800007A
; Patent No. 5834228
;  GENERAL INFORMATION:
;    APPLICANT:  Becker, Joseph
;    APPLICANT:  Nicholson, Donald
;    APPLICANT:  Rotonda, Jennifer
;    APPLICANT:  Thornberry, Nancy
;    APPLICANT:  Fazil, Kimberly
;    APPLICANT:  Gallant, Michel
;    APPLICANT:  Gareau, Yves
;    APPLICANT:  Labelle, Marc
;    APPLICANT:  Peterson, Erin
;    APPLICANT:  Rasper, Dita
;    TITLE OF INVENTION:  CRYSTAL STRUCTURE OF APOPAIN
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  JOSEPH A. COPPOLA - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/800,007A
;      FILING DATE:  13-FEB-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COPPOLA, JOSEPH A
;      REGISTRATION NUMBER:  38,413
;      REFERENCE/DOCKET NUMBER:  19644
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-6734
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  51 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-800-007A-4


  Query Match             2.7%;  Score 31;  DB 2;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-06;
  Matches   31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1026 GCTCACAAAAGAACTCTATTTTTATCACTAA 1056
           |||||||||||||||||||||||||||||||
Db      51 GCTCACAAAAGAACTCTATTTTTATCACTAA 21


RESULT  13
US-08-800-007A-8/c
; Sequence 8, Application US/08800007A
; Patent No. 5834228
;  GENERAL INFORMATION:
;    APPLICANT:  Becker, Joseph
;    APPLICANT:  Nicholson, Donald
;    APPLICANT:  Rotonda, Jennifer
;    APPLICANT:  Thornberry, Nancy
;    APPLICANT:  Fazil, Kimberly
;    APPLICANT:  Gallant, Michel
;    APPLICANT:  Gareau, Yves
;    APPLICANT:  Labelle, Marc
;    APPLICANT:  Peterson, Erin
;    APPLICANT:  Rasper, Dita
;    TITLE OF INVENTION:  CRYSTAL STRUCTURE OF APOPAIN
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  JOSEPH A. COPPOLA - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/800,007A
;      FILING DATE:  13-FEB-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COPPOLA, JOSEPH A
;      REGISTRATION NUMBER:  38,413
;      REFERENCE/DOCKET NUMBER:  19644
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-6734
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  51 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Other
US-08-800-007A-8


  Query Match             2.7%;  Score 31;  DB 2;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-06;
  Matches   31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1026 GCTCACAAAAGAACTCTATTTTTATCACTAA 1056
           |||||||||||||||||||||||||||||||
Db      51 GCTCACAAAAGAACTCTATTTTTATCACTAA 21


RESULT  14
US-09-208-966-20
; Sequence 20, Application US/09208966
; Patent No. 6221355
; GENERAL INFORMATION:
;  APPLICANT: Dowdy, Steven F.
;  TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
;  FILE REFERENCE: 48881/1742
;  CURRENT APPLICATION NUMBER: US/09/208,966
;  CURRENT FILING DATE: 1998-12-10
;  EARLIER APPLICATION NUMBER: 60/082,402
;  EARLIER FILING DATE: 1998-04-20
;  EARLIER APPLICATION NUMBER: 60/069,012
```

; EARLIER FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 90
; TYPE: DNA
; ORGANISM: human
US-09-208-966-20

Query Match 2.6%; Score 30; DB 4; Length 90;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     310 GGAATATCCCTGGACAACAGTTATAAAATG 339
           ||||||||||||||||||||||||||||||
Db      61 ggaatatccctggacaacagttataaaatg 90
```

RESULT 15
US-09-208-966-32
; Sequence 32, Application US/09208966
; Patent No. 6221355
; GENERAL INFORMATION:
; APPLICANT: Dowdy, Steven F.
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
; FILE REFERENCE: 48881/1742
; CURRENT APPLICATION NUMBER: US/09/208,966
; CURRENT FILING DATE: 1998-12-10
; EARLIER APPLICATION NUMBER: 60/082,402
; EARLIER FILING DATE: 1998-04-20
; EARLIER APPLICATION NUMBER: 60/069,012
; EARLIER FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 90
; TYPE: DNA
; ORGANISM: human
US-09-208-966-32

Query Match 2.6%; Score 30; DB 4; Length 90;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     310 GGAATATCCCTGGACAACAGTTATAAAATG 339
           ||||||||||||||||||||||||||||||
Db      61 ggaatatccctggacaacagttataaaatg 90
```

Search completed: August 16, 2002, 01:43:12
Job time: 6203 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 01:48:42 ; Search time 323.42 Seconds
(without alignments)
6152.693 Million cell updates/sec

Title: US-09-613-508-3
Perfect score: 1159
Sequence: 1 GCACGAGCGGATGGGTGCTA.....GGGGACCTACTCTCATGCTG 1159

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1159	100.0	1159	17	AAT15277	Human interleukin-
	2	1118	96.5	1158	17	AAT33567	Pro-Yama cDNA. Ho
	3	1118	96.5	1158	19	AAV05471	Nucleic acid encod
	4	1004	86.6	2635	22	AAS10503	Human caspase 3 DN
	5	1004	86.6	2635	22	AAD03910	Human caspase-3 cD
	6	834	72.0	834	22	AAD13075	Mammalia caspase-3
	7	834	72.0	2538	21	AAC64043	pYNGCas3p10p35 ins
	8	825	71.2	6436	22	AAD13078	pcDNA3-1XUb-C3 con
	9	825	71.2	6703	22	AAD13079	pcDNA3-2XUb-C3 con
	10	825	71.2	6946	22	AAD13080	pcDNA3-3XUb-C3 con
	11	825	71.2	7189	22	AAD13081	pcDNA3-4XUb-C3 con
	12	809	69.8	1001	19	AAV32608	Mutant human apopa
	13	809	69.8	1001	19	AAV32615	Mutant human apopa
	14	783	67.6	835	20	AAX81219	Nucleotide sequenc
	15	753	65.0	7247	22	AAD13084	pcDNA3-Ub-Met-C3 c
	16	476	41.1	873	20	AAX81214	Nucleotide sequenc
	17	438	37.8	3614	22	AAH26878	Anti-HER2 Fab reve
	18	397	34.3	858	20	AAX81215	Nucleotide sequenc
	19	387	33.4	786	22	AAD03918	Caspase chimera ca
	20	325	28.0	443	21	AAC01186	Human secreted pro
	21	52	4.5	834	22	AAI66510	Pig caspase coding
	22	44	3.8	53	22	AAD13082	Mammalia caspase-3
	23	43	3.7	50	22	AAD13076	Mammalia caspase-3
c	24	41	3.5	48	22	AAD13077	Mammalia caspase-3
	25	32	2.8	53	20	AAV33997	Forward primer for
	26	32	2.8	53	20	AAV33993	Forward primer for
c	27	31	2.7	51	20	AAV34000	Reverse primer for
c	28	31	2.7	51	20	AAV33996	Reverse primer for
	29	30	2.6	90	20	AAX78265	pTAT-A/D-p16 fusio
	30	30	2.6	90	20	AAX78253	TAT-CPP32 fusion p
	31	30	2.6	90	21	AAC63839	CPP32 p17 domain/H
	32	30	2.6	90	21	AAC63851	Caspase-3 p17-forw
	33	30	2.6	90	21	AAA46593	PCR primer used to
	34	30	2.6	90	21	AAA46605	PCR primer used to
c	35	29	2.5	52	20	AAV33998	Reverse primer for
c	36	29	2.5	52	20	AAV33994	Reverse primer for
	37	29	2.5	1177	19	AAV13959	Rat interleukin-1
	38	29	2.5	1177	19	AAV13960	Rat interleukin-1
	39	29	2.5	1297	22	AAS10510	Mouse caspase 3 DN
c	40	28	2.4	37	22	AAD03922	Human caspase-3 sm
	41	28	2.4	47	20	AAV33999	Forward primer for
	42	28	2.4	47	20	AAV33995	Forward primer for
	43	27	2.3	27	22	AAS10506	Human caspase 3 fl
	44	25	2.2	31	17	AAT15280	Human ICE-LAP-4 5'
c	45	25	2.2	51	22	AAL27190	Human SNP oligonuc

ALIGNMENTS

RESULT 1
AAT15277
ID AAT15277 standard; cDNA; 1159 BP.
XX
AC AAT15277;
XX
DT 28-OCT-1996 (first entry)
XX
DE Human interleukin-1-converting enzyme-like apoptosis protease-4 cDNA.
XX
KW ICE-LAP-4; interleukin-1-converting enzyme-like apoptosis protease;
KW enzyme; Alzheimer's disease; Parkinson's disease; septic shock;
KW head injury; rheumatoid arthritis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 115..1137
FT /*tag= a
XX
PN WO9613603-A1.
XX
PD 09-MAY-1996.
XX
PF 06-JUN-1995; 95WO-US07235.
XX
PR 01-NOV-1994; 94US-0334251.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Hastings GA, He W, Hudson PL, Rosen CA;
XX

DR WPI; 1996-239509/24.
DR P-PSDB; AAR95831.
XX
PT Human ICE-LAP-3 and -4 DNA and protein - useful in the diagnosis
PT and treatment of Alzheimer's disease, Parkinson's disease,
PT rheumatoid arthritis, septic shock and head injury
XX
PS Claim 1; Page 44; 67pp; English.
XX
CC This sequence was derived from a cDNA library from human endometrial
CC tumuor and may be used for the recombinant production of the ICE-LAP-
CC 4 protein e.g. in COS cells. The encoded protein may be used
CC therapeutically, e.g. as an antitumor or antiviral agent and to
CC control embryonic development and tissue homeostasis. The protein
CC can also be used to treat immunosuppression disorders, such as AIDS,
CC by targeting virus infected cells for cell death. The DNA may find
CC use in gene therapy applications.
XX
SQ Sequence 1159 BP; 329 A; 220 C; 276 G; 334 T; 0 other;

Query Match 100.0%; Score 1159; DB 17; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 gcacgagcggatgggtgctattgtgaggcggttgtagaagagtttcgtgagtgctcgcag 60

Qy      61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ctcatacctgtggctgtgtatccgtggccacagctggttggcgtcgccttgaaatcccag 120

Qy     121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gccgtgaggagttagcgagccctgctcacactcggcgctctggttttcggtgggtgtgcc 180

Qy     181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ctgcacctgcctcttcccgcattctcattaataaaggtatccatggagaacactgaaaac 240

Qy     241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 tcagtggattcaaaatccattaaaaatttggaaccaaagatcatacatggaagcgaatca 300

Qy     301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 atggactctggaatatccctggacaacagttataaaatggattatcctgagatgggttta 360

Qy     361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 tgtataataattaataataagaattttcataaaagcactggaatgacatctcggtctggt 420

Qy     421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 acagatgtcgatgcagcaaacctcagggaaacattcagaaacttgaaatatgaagtcagg 480

Qy     481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 aataaaaatgatcttacacgtgaagaaattgtggaattgatgcgtgatgtttctaaagaa 540

Qy     541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 gatcacagcaaaaggagcagttttgtttgtgtgcttctgagccatggtgaagaaggaata 600

Qy     601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 atttttggaacaaatggacctgttgacctgaaaaaaataacaaacttttcagaggggat 660

Qy     661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 cgttgtagaagtctaactggaaaacccaaacttttcattattcaggcctgccgtggtaca 720

Qy     721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 gaactggactgtggcattgagacagacagtggtgttgatgatgacatggcgtgtcataaa 780

Qy     781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 ataccagtggaggccgacttcttgtatgcatactccacagcacctggttattattcttgg 840

Qy     841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 cgaaattcaaaggatggctcctggttcatccagtcgctttgtgccatgctgaaacagtat 900

Qy     901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 gccgacaagcttgaatttatgcacattcttacccgggttaaccgaaaggtggcaacagaa 960

Qy     961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 tttgagtccttttccttgacgctacttttcatgcaaagaaacagattccatgtattgtt 1020

Qy    1021 TCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 tccatgctcacaaaagaactctatttttatcactaaagaaatggttggttggtggttttt 1080

Qy    1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 tttagtttgtatgccaagtgagaagatggtatatttgggtactgtatttccctctcattg 1140

Qy    1141 GGGACCTACTCTCATGCTG 1159
          ||||||||||||||||||||
Db    1141 gggacctactctcatgctg 1159
```

RESULT 2
AAT33567
ID AAT33567 standard; cDNA; 1158 BP.
XX
AC AAT33567;
XX
DT 06-DEC-1996 (first entry)
XX
DE Pro-Yama cDNA.
XX
KW Yama; CrmA; apoptosis; Fas receptor; T-lymphocyte; AIDS; HIV;
KW human immunodeficiency virus; gene therapy; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 224..1057
FT /*tag= a
XX
PN WO9625501-A2.
XX
PD 22-AUG-1996.
XX
PF 09-FEB-1996; 96WO-US01882.
XX
PR 01-JUN-1995; 95US-0457731.
PR 13-FEB-1995; 95US-0389812.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Dixit VM;
XX
DR WPI; 1996-393402/39.
DR P-PSDB; AAW00677.
XX
PT New nucleic acid encoding Yama protein or CrmA mutant - useful for
PT modulating apoptosis, maintaining T cell viability in AIDS patients
PT and for drug screening
XX
PS Disclosure; Fig 1; 118pp; English.

```
XX
CC   A cDNA clone (AAT33567) codes for pro-Yama (AAW00677), a zymogen which
CC   upon activation cleaves PARP to an 85 kDa form.  Activated Yama has
CC   the ability to modulate cellular function associated with the Fas
CC   receptor pathway such as Fas-associated apoptosis.  This activity is
CC   inhibited by CrmA (see also AAW00678).  The cDNA was isolated from a
CC   cDNA library prepd. from human umbilical vein endothelial cells
CC   treated with tumour necrosis factor and cycloheximide.  It can be
CC   used to produce Yama in a host cell for use in drug screening.
CC   Yama and anti-Yama protein nucleic acids, or antisense RNA, can be
CC   used to modulate cellular functions such as apoptosis and immune
CC   disorders mediated by the Fas pathway.
XX
SQ   Sequence 1158 BP; 328 A; 220 C; 275 G; 335 T; 0 other;

  Query Match             96.5%;  Score 1118;  DB 17;  Length 1158;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1118;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 gcacgagcggatgggtgctattgtgaggcggttgtagaagagtttcgtgagtgctcgcag 61

Qy      61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 ctcatacctgtggctgtgtatccgtggccacagctggttggcgtcgccttgaaatcccag 121

Qy     121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 gccgtgaggagttagcgagccctgctcacactcggcgctctggttttcggtgggtgtgcc 181

Qy     181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 ctgcacctgcctcttcccgcattctcattaataaaggtatccatggagaacactgaaaac 241

Qy     241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 tcagtggattcaaaatccattaaaaatttggaaccaaagatcatacatggaagcgaatca 301

Qy     301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 atggactctggaatatccctggacaacagttataaaatggattatcctgagatgggttta 361

Qy     361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 tgtataataattaataataagaattttcataaaagcactggaatgacatctcggtctggt 421

Qy     421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     422 acagatgtcgatgcagcaaacctcagggaaacattcagaaacttgaaatatgaagtcagg 481

Qy     481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     482 aataaaaatgatcttacacgtgaagaaattgtggaattgatgcgtgatgtttctaaagaa 541

Qy     541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     542 gatcacagcaaaaggagcagttttgtttgtgtgcttctgagccatggtgaagaaggaata 601

Qy     601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     602 atttttggaacaaatggacctgttgacctgaaaaaaataacaaacttttcagaggggat 661

Qy     661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     662 cgttgtagaagtctaactggaaaacccaaacttttcattattcaggcctgccgtggtaca 721

Qy     721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     722 gaactggactgtggcattgagacagacagtggtgttgatgatgacatggcgtgtcataaa 781

Qy     781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     782 ataccagtggaggccgacttcttgtatgcatactccacagcacctggttattattcttgg 841

Qy     841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     842 cgaaattcaaaggatggctcctggttcatccagtcgctttgtgccatgctgaaacagtat 901

Qy     901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     902 gccgacaagcttgaatttatgcacattcttacccgggttaaccgaaaggtggcaacagaa 961

Qy     961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     962 tttgagtccttttccttttgacgctacttttcatgcaaagaaacagattccatgtattgtt 1021

Qy    1021 TCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1022 tccatgctcacaaaagaactctatttttatcactaaagaaatggttggttggtggttttt 1081

Qy    1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGG 1118
           ||||||||||||||||||||||||||||||||||||
Db    1082 tttagtttgtatgccaagtgagaagatggtatatttgg 1119


RESULT   3
AAV05471
ID   AAV05471 standard; DNA; 1158 BP.
XX
AC   AAV05471;
XX
DT   02-JUL-1998  (first entry)
XX
DE   Nucleic acid encoding a protein designated YAMA.
XX
KW   CrmA; tumour necrosis factor receptor; TNF-R; U1-70; inhibition;
KW   apoptosis; activation; U1-70 path; apoptosis path; identification;
KW   modulation; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             224..1057
FT                   /*tag=  a
XX
PN   JP09299077-A.
XX
PD   25-NOV-1997.
XX
PF   27-JAN-1997;   97JP-0012932.
XX
PR   09-FEB-1996;   96US-0591605.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
DR   WPI; 1998-056551/06.
DR   P-PSDB; AAW41688.
XX
PT   Controlling apoptosis by inhibiting activation of the U1-70 pathway
PT   - by introduction of nucleic acid encoding protein with CrmA
PT   biological activity into the cell
XX
PS   Disclosure; Fig 1; 48pp; Japanese.
XX
CC   The present sequence encodes a non-naturally occurring protein designated
CC   YAMA. The specification describes a novel method of controlling cell
CC   functions controlled by the tumour necrosis factor receptor (TNF-R)
CC   pathway containing U1-70. The method comprises introducing a nucleic
CC   acid molecule encoding CrmA into a cell, and culturing the cell under
CC   suitable conditions for transcription and translation of the CrmA nucleic
CC   acid. Method for preventing and inhibiting apoptosis by inhibiting
CC   activation of the U1-70 path, identifying the chemical related to the
CC   apoptosis path, and screening a candidate chemical having biological
CC   function in the apoptosis path are also described. The methods can be
```

```
CC   used to identify substances which modulate apoptosis in cells controlled
CC   by the TNF-R pathway.
XX
SQ   Sequence 1158 BP; 328 A; 220 C; 275 G; 335 T; 0 other;


  Query Match             96.5%;  Score 1118;  DB 19;  Length 1158;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1118;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 gcacgagcggatgggtgctattgtgaggcggttgtagaagagtttcgtgagtgctcgcag 61

Qy      61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 ctcatacctgtggctgtgtatccgtggccacagctggttggcgtcgccttgaaatcccag 121

Qy     121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 gccgtgaggagttagcgagccctgctcacactcggcgctctggttttcggtgggtgtgcc 181

Qy     181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 ctgcacctgcctcttcccgcattctcattaataaaggtatccatggagaacactgaaaac 241

Qy     241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 tcagtggattcaaaatccattaaaaatttggaaccaaagatcatacatggaagcgaatca 301

Qy     301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 atggactctggaatatccctggacaacagttataaaatggattatcctgagatgggttta 361

Qy     361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 tgtataataattaataataagaattttcataaaagcactggaatgacatctcggtctggt 421

Qy     421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     422 acagatgtcgatgcagcaaacctcagggaaacattcagaaacttgaaatatgaagtcagg 481

Qy     481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     482 ataaaaatgatcttacacgtgaagaaattgtggaattgatgcgtgatgtttctaaagaa 541

Qy     541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     542 gatcacagcaaaaggagcagttttgtttgtgtgcttctgagccatggtgaagaaggaata 601

Qy     601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     602 atttttggaacaaatggacctgttgacctgaaaaaaataacaaacttttttcagaggggat 661

Qy     661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     662 cgttgtagaagtctaactggaaaacccaaacttttcattattcaggcctgccgtggtaca 721

Qy     721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     722 gaactggactgtggcattgagacagacagtggtgttgatgatgacatggcgtgtcataaa 781

Qy     781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     782 ataccagtggaggccgacttcttgtatgcatactccacagcacctggttattattcttgg 841

Qy     841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     842 cgaaattcaaaggatggctcctggttcatccagtcgctttgtgccatgctgaaacagtat 901

Qy     901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     902 gccgacaagcttgaatttatgcacattcttacccgggttaaccgaaaggtggcaacagaa 961

Qy     961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     962 tttgagtccttttccttttgacgctacttttcatgcaaagaaacagattccatgtattgtt 1021

Qy    1021 TCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1022 tccatgctcacaaaagaactctatttttatcactaaagaaatggttggttggtggttttt 1081

Qy    1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGG 1118
          ||||||||||||||||||||||||||||||||||||||
Db    1082 tttagtttgtatgccaagtgagaagatggtatatttgg 1119


RESULT   4
AAS10503
ID   AAS10503 standard; DNA; 2635 BP.
XX
AC   AAS10503;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Human caspase 3 DNA sequence.
XX
KW   Human; caspase 3; apoptosis; hyperproliferative disorder; hepatitis;
KW   viral infection; haematopoietic disorder; autoimmune disorder;
KW   atherosclerosis; neurological disorder; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             225..1058
FT                   /*tag= a
FT                   /product= "Caspase 3"
XX
PN   WO200153310-A1.
XX
PD   26-JUL-2001.
XX
PF   11-JAN-2001; 2001WO-US00888.
XX
PR   18-JAN-2000; 2000US-0484617.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Zhang H,  Cowsert LM;
XX
DR   WPI; 2001-442252/47.
DR   P-PSDB; AAU05394.
XX
PT   New antisense compound to inhibit caspase 3 is useful for treating
PT   hepatitis and atherosclerosis -
XX
PS   Example 13; Page 92-94; 127pp; English.
XX
CC   The present sequence encoding for human caspase 3 is described
CC   in an invention relating to novel antisense oligonucleotides
CC   (AAS10517-AAS10676) and methods of using these compounds for the
CC   modulation of caspase 3 expression. The caspase 3 antisense
CC   oligonucleotides specifically hybridise with and inhibit the
CC   expression of caspase 3. Antisense compounds targeted to caspase 3
CC   are useful to inhibit caspase 3 expression in cells or tissues and
CC   to modulate apoptosis. The caspase 3 antisense oligonucleotides are
CC   useful for treating disorders associated with expression of caspase 3.
CC   Such disorders include hyperproliferative disorders (e.g. cancer),
CC   viral infections (e.g. hepatitis), haematopoietic disorders, autoimmune
CC   disorders, atherosclerosis and neurological disorders (e.g. Alzheimer's
CC   disease).
XX
SQ   Sequence 2635 BP; 815 A; 465 C; 569 G; 786 T; 0 other;

  Query Match             86.6%;  Score 1004;  DB 22;  Length 2635;
```

```
 Best Local Similarity   99.8%;  Pred. No. 0;
 Matches 1104;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      13 GGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTG 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 gggtgctattgtgaggcggttgtagaagagtttcgtgagtgctcgcagctcatacctgtg 74

Qy      73 GCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGT 132
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 gctgtgtatccgtggccacagctggttggcgtcgccttgaaatcccaggccgtgaggagt 134

Qy     133 TAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 tagcgagccctgctcacactcggcgctctggttttcggtgggtgtgccctgcacctgcct 194

Qy     193 CTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCA 252
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     195 cttcccccattctcattaataaaggtatccatggagaacactgaaaactcagtggattca 254

Qy     253 AAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGA 312
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     255 aaatccattaaaaatttggaaccaaagatcatacatggaagcgaatcaatggactctgga 314

Qy     313 ATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     315 atatccctggacaacagttataaaatggattatcctgagatgggtttatgtataataatt 374

Qy     373 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGAT 432
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     375 aataataagaattttcataaaagcactggaatgacatctcggtctggtacagatgtcgat 434

Qy     433 GCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGAT 492
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     435 gcagcaaacctcagggaaacattcagaaacttgaaatatgaagtcaggaataaaaatgat 494

Qy     493 CTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAA 552
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     495 cttacacgtgaagaaattgtggaattgatgcgtgatgtttctaaagaagatcacagcaaa 554

Qy     553 AGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACA 612
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     555 aggagcagttttgtttgtgtgcttctgagccatggtgaagaaggaataatttttggaaca 614

Qy     613 AATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 672
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     615 aatggacctgttgacctgaaaaaaataacaaacttttcagaggggatcgttgtagaagt 674

Qy     673 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGT 732
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     675 ctaactggaaaacccaaacttttcattattcaggcctgccgtggtacagaactggactgt 734

Qy     733 GGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAG 792
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     735 ggcattgagacagacagtggtgttgatgatgacatggcgtgtcataaaataccagtggat 794

Qy     793 GCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAG 852
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     795 gccgacttcttgtatgcatactccacagcacctggttattattcttggcgaaattcaaag 854

Qy     853 GATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTT 912
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     855 gatggctcctggttcatccagtcgctttgtgccatgctgaaacagtatgccgacaagctt 914

Qy     913 GAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 972
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     915 gaatttatgcacattcttacccgggttaaccgaaaggtggcaacagaatttgagtccttt 974

Qy     973 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACA 1032
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     975 tcctttgacgctacttttcatgcaaagaaacagattccatgtattgtttccatgctcaca 1034

Qy    1033 AAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTAT 1092
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    1035 aaagaactctatttttatcactaaagaaatggttggttggtggttttttttagtttgtat 1094

Qy    1093 GCCAAGTGAGAAGATGGTATATTTGG 1118
           |||||||||||||||||||||||||
Db    1095 gccaagtgagaagatggtatatttgg 1120


RESULT   5
AAD03910
ID   AAD03910 standard; cDNA; 2635 BP.
XX
AC   AAD03910;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Human caspase-3 cDNA.
XX
KW   Human; caspase-3; CPP32; yama; apopain;
KW   cysteine protease; apoptosis; caspase expression cassette; metastasis;
KW   tumour; cathepsin B; urokinase; proliferation; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             225..1058
FT                   /*tag= a
FT                   /product= "Human caspase-3"
XX
PN   WO200129232-A2.
XX
PD   26-APR-2001.
XX
PF   19-OCT-2000; 2000WO-US28941.
XX
PR   20-OCT-1999;   99US-0160559.
PR   14-AUG-2000; 2000US-0225564.
XX
PA   (SCIO-) SCIOS INC.
XX
PI   Cordell B,  Li Y;
XX
DR   WPI; 2001-290920/30.
DR   P-PSDB; AAE00600.
XX
PT   Novel fusion polypeptide comprising first and second caspase subunit
PT   separated by cleavage site not associated in nature with caspase
PT   subunit, useful for cloning gene encoding enzymes involved in
PT   proteolytic cleavage
XX
PS   Example 1; Fig 5; 116pp; English.
XX
CC   The present sequence is a cDNA encoding human Caspase-3 also known as
CC   CPP32, yama and apopain. Caspases are a family of cysteine proteases,
CC   that participate in the initiation and execution of apoptosis. Caspases
CC   exist as pro-enzymes, activated by cleavage into a large and small
CC   subunit, occurring after specific aspartic acid residues within the
CC   pro-enzyme sequence. The present invention relates to a method for
CC   functional cloning of genes encoding proteins or enzymes involved in
CC   proteolytic cleavage. The invention is based on the use of caspase
CC   expression cassettes comprising the coding sequence of a proteolytic
CC   cleavage site flanked by sequences encoding two caspase subunits. A
CC   fusion polypeptide comprising a first and a second caspase subunit,
CC   separated by a cleavage site not associated in nature, is useful for
CC   cloning gene encoding enzymes involved in proteolytic cleavage. An
CC   expression cassette containing fusion polypeptide is used to identify a
CC   mutant cell line deficient in an enzyme of interest and is also useful
CC   for diagnosis and suppression of proliferation or metastases of a tumour
CC   cell characterised by overexpression of a polypeptide (e.g. Cathepsin B
CC   or urokinase, selectively expressed in the tumour cells). DNA encoding
CC   fusion polypeptide is used in gene therapy.
XX
SQ   Sequence 2635 BP; 815 A; 465 C; 569 G; 786 T; 0 other;
```

```
Query Match             86.6%;  Score 1004;  DB 22;  Length 2635;
Best Local Similarity   99.8%;  Pred. No. 0;
Matches 1104;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      13 GGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTG 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 gggtgctattgtgaggcggttgtagaagagtttcgtgagtgctcgcagctcatacctgtg 74

Qy      73 GCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGT 132
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 gctgtgtatccgtggccacagctggttggcgtcgccttgaaatcccaggccgtgaggagt 134

Qy     133 TAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 tagcgagccctgctcacactcggcgctctggttttcggtgggtgtgccctgcacctgcct 194

Qy     193 CTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCA 252
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     195 cttcccccattctcattaataaaggtatccatggagaacactgaaaactcagtggattca 254

Qy     253 AAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGA 312
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     255 aaatccattaaaaatttggaaccaaagatcatacatggaagcgaatcaatggactctgga 314

Qy     313 ATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     315 atatccctggacaacagttataaaatggattatcctgagatgggtttatgtataataatt 374

Qy     373 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGAT 432
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     375 aataataagaattttcataaaagcactggaatgacatctcggtctggtacagatgtcgat 434

Qy     433 GCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGAT 492
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     435 gcagcaaacctcagggaaacattcagaaacttgaaatatgaagtcaggaataaaaatgat 494

Qy     493 CTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAA 552
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     495 cttacacgtgaagaaattgtggaattgatgcgtgatgtttctaaagaagatcacagcaaa 554

Qy     553 AGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACA 612
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     555 aggagcagttttgtttgtgtgcttctgagccatggtgaagaaggaataatttttggaaca 614

Qy     613 AATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 672
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     615 aatggacctgttgacctgaaaaaaataacaaacttttcagaggggatcgttgtagaagt 674

Qy     673 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGT 732
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     675 ctaactggaaaacccaaacttttcattattcaggcctgccgtggtacagaactggactgt 734

Qy     733 GGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAG 792
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     735 ggcattgagacagacagtggtgttgatgatgacatggcgtgtcataaaatacCagtggat 794

Qy     793 GCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAG 852
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     795 gccgacttcttgtatgcatactccacagcacctggttattattcttggcgaaattcaaag 854

Qy     853 GATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTT 912
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     855 gatggctcctggttcatccagtcgctttgtgccatgctgaaacagtatgccgacaagctt 914

Qy     913 GAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 972
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     915 gaatttatgcacattcttacccgggttaaccgaaaggtggcaacagaatttgagtccttt 974

Qy     973 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACA 1032
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     975 tcctttgacgctacttttcatgcaaagaaacagattccatgtattgtttccatgctcaca 1034

Qy     1033 AAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTAT 1092
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1035 aaagaactctatttttatcactaaagaaatggttggttggtggttttttttagtttgtat 1094

Qy     1093 GCCAAGTGAGAAGATGGTATATTTGG 1118
            |||||||||||||||||||||||||
Db     1095 gccaagtgagaagatggtatatttgg 1120
```

RESULT 6
AAD13075
ID AAD13075 standard; cDNA; 834 BP.
XX
AC AAD13075;
XX
DT 16-OCT-2001 (first entry)
XX
DE Mammalia caspase-3 reporter moiety cDNA.
XX
KW Protein destabilisation; linker moiety; reporter moiety; disease model;
KW linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW transgenic animal; transgenic plant; disease resistance;
KW caspase-3; ss.
XX
OS Mammalia.
XX
PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US03791.
XX
PR 04-FEB-2000; 2000US-0498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by
PT -NH-ubiquitin protein endoproteases, comprising two copies of the
PT domain -
XX
PS Example 4; Page 131; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method
CC is also useful for developing novel assays for a wide range of
CC post-translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is mammalia caspase-3 cDNA which is used as a reporter moiety in
CC the exemplification of the invention.
XX
SQ Sequence 834 BP; 274 A; 147 C; 179 G; 234 T; 0 other;

```
Query Match             72.0%;  Score 834;  DB 22;  Length 834;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches  834;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     223 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATC 282
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1 atggagaacactgaaaactcagtggattcaaaatccattaaaaatttggaaccaaagatc 60

Qy    283 ATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGAT 342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 atacatggaagcgaatcaatggactctggaatatccctggacaacagttataaaatggat 120

Qy    343 TATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGA 402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 tatcctgagatgggtttatgtataataattaataataagaattttcataaaagcactgga 180

Qy    403 ATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAAC 462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 atgacatctcggtctggtacagatgtcgatgcagcaaacctcagggaaacattcagaaac 240

Qy    463 TTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaaattgtggaattgatg 300

Qy    523 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGC 582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 cgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtttgtgtgcttctgagc 360

Qy    583 CATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACA 642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 catggtgaagaaggaataatttttggaacaaatggacctgttgacctgaaaaaaataaca 420

Qy    643 AACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATT 702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 aacttttcagaggggatcgttgtagaagtctaactggaaaacccaaacttttcattatt 480

Qy    703 CAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGAT 762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 caggcctgccgtggtacagaactggactgtggcattgagacagacagtggtgttgatgat 540

Qy    763 GACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 gacatggcgtgtcataaaatacCagtggaggccgacttcttgtatgcatactccacagca 600

Qy    823 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGT 882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 cctggttattattcttggcgaaattcaaaggatggctcctggttcatccagtcgctttgt 660

Qy    883 GCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAAC 942
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 gccatgctgaaacagtatgccgacaagcttgaatttatgcacattcttacccgggttaac 720

Qy    943 CGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAA 1002
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 cgaaaggtggcaacagaatttgagtccttttccttgacgctacttttcatgcaaagaaa 780

Qy   1003 CAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAA 1056
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 cagattccatgtattgtttccatgctcacaaaagaactctatttttatcactaa 834


RESULT   7
AAC64043
ID   AAC64043 standard; DNA; 2538 BP.
XX
AC   AAC64043;
XX
DT   14-FEB-2001  (first entry)
XX
DE   pYNGCas3p10p35 insert encoding human caspase-3 and AcNPV p35 protein.
XX
KW   Human caspase-3; AcNPV p35 protein; BmNPV p10 promoter;
KW   polyhedron promoter; baculovirus vector; apoptosis inhibitor protein;
KW   recominant caspase production; insect cell; ds.
XX
OS   Chimeric - Homo sapiens.
OS   Chimeric - Autographa californica nuclear polyhedrosis virus.
OS   Chimeric - Bombyx mori nuclear polyhedrosis virus.
XX
PN   JP2000253882-A.
XX
PD   19-SEP-2000.
XX
PF   10-MAR-1999;   99JP-0062732.
XX
PR   10-MAR-1999;   99JP-0062732.
XX
PA   (KATA ) KATAKURA IND CO LTD.
XX
DR   WPI; 2000-658811/64.
XX
PT   Baculovirus vector used for the preparation of caspase protein,
PT   comprises a transfer vector plasmid containing a caspase
PT   protein-producing gene -
XX
PS   Example 1; Page 7-8; 10pp; Japanese.
XX
CC   The invention relates to a novel baculovirus vector used for the
CC   preparation of a caspase protein, and containing a recombinant transfer
CC   vector plasmid comprising a gene encoding a caspase and a gene encoding
CC   an apoptosis-inhibiting protein. The invention also encompasses the
CC   preparation of caspase protein, comprising transfecting the novel
CC   baculovirus vector into an insect or an insect cell; breeding the
CC   insect or culturing the cell; and collecting the caspase protein from
CC   the insect or insect cell. The present sequence represents the expression
CC   cassette insert from a transfer vector (pYNGCas3p10p35) produced in an
CC   exemplification of the invention. This insert comprises a mature human
CC   caspase-3 gene under the control of a polyhedron promoter, and an
CC   Autographa californica nuclear polyhedrosis virus (AcNPV) p35 gene under
CC   the control of a Bombyx mori nuclear polyhedrosis virus (BmNPV) p10
CC   promoter.
XX
SQ   Sequence 2538 BP; 865 A; 446 C; 506 G; 721 T; 0 other;


  Query Match             72.0%;  Score 834;  DB 21;  Length 2538;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  834;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    223 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATC 282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    136 atggagaacactgaaaactcagtggattcaaaatccattaaaaatttggaaccaaagatc 195

Qy    283 ATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGAT 342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    196 atacatggaagcgaatcaatggactctggaatatccctggacaacagttataaaatggat 255

Qy    343 TATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGA 402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    256 tatcctgagatgggtttatgtataataattaataataagaattttcataaaagcactgga 315

Qy    403 ATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAAC 462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    316 atgacatctcggtctggtacagatgtcgatgcagcaaacctcagggaaacattcagaaac 375

Qy    463 TTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    376 ttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaaattgtggaattgatg 435

Qy    523 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGC 582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    436 cgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtttgtgtgcttctgagc 495

Qy    583 CATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACA 642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    496 catggtgaagaaggaataatttttggaacaaatggacctgttgacctgaaaaaaataaca 555

Qy    643 AACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATT 702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    556 aacttttcagaggggatcgttgtagaagtctaactggaaaacccaaacttttcattatt 615

Qy    703 CAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGAT 762
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     616 caggcctgccgtggtacagaactggactgtggcattgagacagacagtggtgttgatgat 675

Qy     763 GACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 822
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     676 gacatggcgtgtcataaaatacCagtggaggccgacttcttgtatgcatactccacagca 735

Qy     823 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGT 882
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     736 cctggttattattcttggcgaaattcaaaggatggctcctggttcatccagtcgctttgt 795

Qy     883 GCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAAC 942
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     796 gccatgctgaaacagtatgccgacaagcttgaatttatgcacattcttacccgggttaac 855

Qy     943 CGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAA 1002
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     856 cgaaaggtggcaacagaatttgagtccttttccttgacgctacttttcatgcaaagaaa 915

Qy    1003 CAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAA 1056
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     916 cagattccatgtattgtttccatgctcacaaaagaactctatttttatcactaa 969


RESULT   8
AAD13078
ID   AAD13078 standard; DNA; 6436 BP.
XX
AC   AAD13078;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-1XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-1XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain  -
XX
PS   Example 4; Page 133-134; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-1XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domain-Mammalia caspase-3 fusion protein.
XX
SQ   Sequence 6436 BP; 1577 A; 1606 C; 1607 G; 1646 T; 0 other;


  Query Match             71.2%;  Score 825;  DB 22;  Length 6436;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  825;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     229 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1144 aacactgaaaactcagtggattcaaaatccattaaaaatttggaaccaaagatcatacat 1203

Qy     289 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 348
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1204 ggaagcgaatcaatggactctggaatatccctggacaacagttataaaatggattatcct 1263

Qy     349 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 408
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1264 gagatgggtttatgtataataattaataataagaattttcataaaagcactggaatgaca 1323

Qy     409 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 468
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1324 tctcggtctggtacagatgtcgatgcagcaaacctcagggaaacattcagaaacttgaaa 1383

Qy     469 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 528
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1384 tatgaagtcaggaataaaaatgatcttacacgtgaagaaattgtggaattgatgcgtgat 1443

Qy     529 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 588
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1444 gtttctaaagaagatcacagcaaaaggagcagttttgtttgtgtgcttctgagccatggt 1503

Qy     589 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 648
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1504 gaagaaggaataatttttggaacaaatggacctgttgacctgaaaaaaataacaaacttt 1563

Qy     649 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 708
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1564 ttcagaggggatcgttgtagaagtctaactggaaaacccaaacttttcattattcaggcc 1623

Qy     709 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 768
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1624 tgccgtggtacagaactggactgtggcattgagacagacagtggtgttgatgatgacatg 1683

Qy     769 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 828
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1684 gcgtgtcataaaatacCagtggaggccgacttcttgtatgcatactccacagcacctggt 1743

Qy     829 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 888
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1744 tattattcttggcgaaattcaaaggatggctcctggttcatccagtcgctttgtgccatg 1803

Qy     889 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1804 ctgaaacagtatgccgacaagcttgaatttatgcacattcttacccgggttaaccgaaag 1863

Qy     949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1008
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1864 gtggcaacagaatttgagtccttttccttgacgctacttttcatgcaaagaaacagatt 1923

Qy    1009 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1053
           |||||||||||||||||||||||||||||||||||||||||||||
Db    1924 ccatgtattgtttccatgctcacaaaagaactctatttttatcac 1968


RESULT   9
```

```
AAD13079
ID   AAD13079 standard; DNA; 6703 BP.
XX
AC   AAD13079;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-2XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-2XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain -
XX
PS   Example 4; Page 134-135; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-2XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domains-Mammalia caspase-3
CC   fusion protein.
XX
SQ   Sequence 6703 BP; 1653 A; 1677 C; 1674 G; 1699 T; 0 other;


  Query Match             71.2%;  Score 825;  DB 22;  Length 6703;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  825;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     229 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1411 aacactgaaaactcagtggattcaaaatccattaaaaatttggaaccaaagatcatacat 1470

Qy     289 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 348
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1471 ggaagcgaatcaatggactctggaatatccctggacaacagttataaaatggattatcct 1530

Qy     349 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 408
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1531 gagatgggtttatgtataataattaataataagaattttcataaaagcactggaatgaca 1590

Qy     409 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 468
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1591 tctcggtctggtacagatgtcgatgcagcaaacctcagggaaacattcagaaacttgaaa 1650

Qy     469 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 528
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1651 tatgaagtcaggaataaaaatgatcttacacgtgaagaaattgtggaattgatgcgtgat 1710

Qy     529 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 588
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1711 gtttctaaagaagatcacagcaaaaggagcagttttgtttgtgtgcttctgagccatggt 1770

Qy     589 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 648
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1771 gaagaaggaataatttttggaacaaatggacctgttgacctgaaaaaaataacaaacttt 1830

Qy     649 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 708
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1831 ttcagaggggatcgttgtagaagtctaactggaaaacccaaacttttcattattcaggcc 1890

Qy     709 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 768
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1891 tgccgtggtacagaactggactgtggcattgagacagacagtggtgttgatgatgacatg 1950

Qy     769 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 828
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1951 gcgtgtcataaaatacca gtggaggccgacttcttgtatgcatactccacagcacctggt 2010

Qy     829 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 888
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2011 tattattcttggcgaaattcaaaggatggctcctggttcatccagtcgctttgtgccatg 2070

Qy     889 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2071 ctgaaacagtatgccgacaagcttgaatttatgcacattcttacccgggttaaccgaaag 2130

Qy     949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1008
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2131 gtggcaacagaatttgagtccttttccttt gacgctacttttcatgcaaagaaacagatt 2190

Qy    1009 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1053
           |||||||||||||||||||||||||||||||||||||||||||||
Db    2191 ccatgtattgtttccatgctcacaaaagaactctatttttatcac 2235


RESULT  10
AAD13080
ID   AAD13080 standard; DNA; 6946 BP.
XX
AC   AAD13080;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-3XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-3XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
```

PR 04-FEB-2000; 2000US-0498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by
PT -NH-ubiquitin protein endoproteases, comprising two copies of the
PT domain -
XX
PS Example 4; Page 136-137; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method
CC is also useful for developing novel assays for a wide range of
CC post-translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a pcDNA3-3XUb-C3 construct comprising mutant multimerised
CC human ubiquitin destabilisation domains-Mammalia caspase-3
CC fusion protein.
XX
SQ Sequence 6946 BP; 1723 A; 1742 C; 1735 G; 1746 T; 0 other;

Query Match 71.2%; Score 825; DB 22; Length 6946;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 825; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    229 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 288
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1654 aacactgaaaactcagtggattcaaaatccattaaaaatttggaaccaaagatcatacat 1713

Qy    289 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 348
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1714 ggaagcgaatcaatggactctggaatatccctggacaacagttataaaatggattatcct 1773

Qy    349 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 408
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1774 gagatgggtttatgtataataattaataataagaattttcataaaagcactggaatgaca 1833

Qy    409 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 468
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1834 tctcggtctggtacagatgtcgatgcagcaaacctcagggaaacattcagaaacttgaaa 1893

Qy    469 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 528
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1894 tatgaagtcaggaataaaaatgatcttacacgtgaagaaattgtggaattgatgcgtgat 1953

Qy    529 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 588
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1954 gtttctaaagaagatcacagcaaaaggagcagttttgtttgtgtgcttctgagccatggt 2013

Qy    589 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 648
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2014 gaagaaggaataatttttggaacaaatggacctgttgacctgaaaaaaataacaaacttt 2073

Qy    649 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 708
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2074 ttcagaggggatcgttgtagaagtctaactggaaaacccaaacttttcattattcaggcc 2133

Qy    709 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 768
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2134 tgccgtggtacagaactggactgtggcattgagacagacagtggtgttgatgatgacatg 2193

Qy    769 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 828
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2194 gcgtgtcataaaatacCagtggaggccgacttcttgtatgcatactccacagcacctggt 2253

Qy    829 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 888
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2254 tattattcttggcgaaattcaaaggatggctcctggttcatccagtcgctttgtgccatg 2313

Qy    889 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2314 ctgaaacagtatgccgacaagcttgaatttatgcacattcttacccgggttaaccgaaag 2373

Qy    949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1008
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2374 gtggcaacagaatttgagtcctttTccTttgacgctacttttcatgcaaagaaacagatt 2433

Qy    1009 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1053
           |||||||||||||||||||||||||||||||||||||||||||||
Db   2434 ccatgtattgtttccatgctcacaaaagaactctatttttatcac 2478
```

RESULT 11
AAD13081
ID AAD13081 standard; DNA; 7189 BP.
XX
AC AAD13081;
XX
DT 16-OCT-2001 (first entry)
XX
DE pcDNA3-4XUb-C3 construct.
XX
KW Protein destabilisation; linker moiety; reporter moiety; disease model;
KW linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW transgenic animal; transgenic plant; disease resistance; caspase-3;
KW human; ubiquitin; cyclic; circular; plasmid pcDNA3-4XUb-C3;
KW fusion protein; ds.
XX
OS Chimeric - Mammalia.
OS Chimeric - Homo sapiens.
XX
PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US03791.
XX
PR 04-FEB-2000; 2000US-0498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by
PT -NH-ubiquitin protein endoproteases, comprising two copies of the
PT domain -
XX
PS Example 4; Page 137-138; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method
CC is also useful for developing novel assays for a wide range of
CC post-translational activities, such as proteolysis, phosphorylation,

```
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-4XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domains-Mammalia caspase-3
CC   fusion protein.
XX
SQ   Sequence 7189 BP; 1793 A; 1807 C; 1796 G; 1793 T; 0 other;


  Query Match           71.2%;  Score 825;  DB 22;  Length 7189;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  825;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     229 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1897 aacactgaaaactcagtggattcaaaatccattaaaaatttggaaccaaagatcatacat 1956

Qy     289 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 348
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1957 ggaagcgaatcaatggactctggaatatccctggacaacagttataaaatggattatcct 2016

Qy     349 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 408
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2017 gagatgggtttatgtataatAattaataataagaattttcataaaagcactggaatgaca 2076

Qy     409 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 468
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2077 tctcggtctggtacagatgtcgatgcagcaaacctcagggaaacattcagaaacttgaaa 2136

Qy     469 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 528
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2137 tatgaagtcaggaataaaaatgatcttacacgtgaagaaattgtggaattgatgcgtgat 2196

Qy     529 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 588
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2197 gtttctaaagaagatcacagcaaaaggagcagttttgtttgtgtgcttctgagccatggt 2256

Qy     589 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 648
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2257 gaagaaggaataatttttggaacaaatggacctgttgacctgaaaaaaataacaaacttt 2316

Qy     649 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 708
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2317 ttcagaggggatcgttgtagaagtctaactggaaaacccaaacttttcattattcaggcc 2376

Qy     709 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 768
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2377 tgccgtggtacagaactggactgtggcattgagacagacagtggtgttgatgatgacatg 2436

Qy     769 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 828
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2437 gcgtgtcataaaatACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2496

Qy     829 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 888
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2497 tattattcttggcgaaattcaaaggatggctcctggttcatccagtcgctttgtgccatg 2556

Qy     889 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2557 ctgaaacagtatgccgacaagcttgaatttatgcacattcttacccgggttaaccgaaag 2616

Qy     949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1008
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2617 gtggcaacagaatttgagtcctttttcctttgacgctacttttcatgcaaagaaacagatt 2676

Qy    1009 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1053
           |||||||||||||||||||||||||||||||||||||||||||||
Db    2677 ccatgtattgtttccatgctcacaaaagaactctatttttatcac 2721
```

```
RESULT  12
AAV32608
ID   AAV32608 standard; cDNA; 1001 BP.
XX
AC   AAV32608;
XX
DT   26-OCT-1998  (first entry)
XX
DE   Mutant human apopain (caspase-3) pro-enzyme cDNA.
XX
KW   Mutant human apopain pro-enzyme; caspase-3; protease; SPA;
KW   scintillation proximity binding assay; diabetes; cancer;
KW   4-phosphono(difluoromethyl) phenylalanine group; osteoporosis;
KW   tyrosine phosphatase; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             64..897
FT                   /*tag= a
FT                   /product= "Mutant apopain C163S pro-enzyme"
FT   misc_feature    148..588
FT                   /*tag= b
FT                   /note= "Encodes the P17 subunit"
FT   mutation        550..552
FT                   /*tag= c
FT                   /note= "Changed from TGC in wild-type to TCC in
FT                   mutant"
FT   misc_feature    589..894
FT                   /*tag= d
FT                   /note= "Encodes the P12 subunit"
XX
PN   WO9820024-A1.
XX
PD   14-MAY-1998.
XX
PF   03-NOV-1997;   97WO-CA00824.
XX
PR   04-NOV-1996;   96US-0030411.
XX
PA   (MERI ) MERCK FROSST CANADA INC.
XX
PI   Desmarais S,  Friesen R,  Zamboni R;
XX
DR   WPI; 1998-348101/30.
DR   P-PSDB; AAW48937.
XX
PT   Peptide(s) useful in binding assays for tyrosine phosphatases or
PT   cysteine proteases - contain two or more 4-phosphono(difluoromethyl)
PT   phenylalanine groups to improve binding affinity
XX
PS   Disclosure; Pages 38-39; 59pp; English.
XX
CC   The present sequence represents a mutated human apopain (caspase-3)
CC   pro-enzyme cDNA.  The cDNA encodes a mutant human apopain C163S
CC   pro-enzyme.  The invention provides a method for use in a
CC   scintillation proximity binding assay (SPA) for proteases and
CC   phosphatases.  The method involves using mutated proteases and
CC   phosphatases whereby the catalytic cysteine residue of the enzymes
CC   are replaced with a serine or alanine residue to correct the problem of
CC   interference in SPA from extraneous oxidising and alkylating agents.
CC   The mutation affects the catalytic properties of the enzyme but does
CC   not affect their binding properties.  The invention claims for new
CC   ligands for use in SPA which have increased binding affinity for a
CC   tyrosine phosphatase or cysteine protease.  The ligands contain at least
CC   two 4-phosphono(difluoromethyl) phenylalanine groups which increase
CC   binding affinity of the ligand to its respective enzyme.  The assay can
CC   be used to determine the ability of new ligands and compound mixtures to
CC   competitively bind with an enzyme.  The method is claimed to allow
CC   a better usage of SPA in the discovery of compounds for the treatment
CC   and study of diseases, e.g. diabetes, cancer and osteoporosis.
XX
```

```
SQ   Sequence 1001 BP; 310 A; 172 C; 226 G; 293 T; 0 other;

  Query Match             69.8%;  Score 809;  DB 19;  Length 1001;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches  909;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49 ttaataaaggtatccatggagaacactgaaaactcagtggattcaaaatccattaaaaat 108

Qy     268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     109 ttggaaccaaagatcatacatggaagcgaatcaatggactctggaatatccctggacaac 168

Qy     328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     169 agttataaaatggattatcctgagatgggtttatgtataataattaataataagaatttt 228

Qy     388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     229 cataagagcactggaatgacatctcggtctggtacagatgtcgatgcagcaaacctcagg 288

Qy     448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     289 gaaacattcagaaacttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaa 348

Qy     508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     349 attgtggaattgatgcgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtt 408

Qy     568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     409 tgtgtgcttctgagccatggtgaagaaggaataatttttggaacaaatggacctgttgac 468

Qy     628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     469 ctgaaaaaaataacaaacttttttcagaggggatcgttgtagaagtctaactggaaaaccc 528

Qy     688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     529 aaacttttcattattcaggcctcccgtggtacagaactggactgtggcattgagacagac 588

Qy     748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     589 agtggtgttgatgatgacatggcgtgtcataaaatacCagtggaggccgacttcttgtat 648

Qy     808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     649 gcatactccacagcacctggttattattcttggcgaaattcaaaggatggctcctggttc 708

Qy     868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     709 atccagtcgctttgtgccatgctgaaacagtatgccgacaagcttgaatttatgcacatt 768

Qy     928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     769 cttacccgggttaaccgaaaggtggcaacagaatttgagtccttttccttgacgctact 828

Qy     988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     829 tttcatgcaaagaaacagattccatgtattgtttccatgctcacaaaagaactctatttt 888

Qy    1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     889 tatcactaaagaaatggttggttggtggttttttttagtttgtatgccaagtgagaagat 948

Qy    1108 GGTATATTTGG 1118
           ||||||||||| 
Db     949 ggtatatttgg 959

RESULT  13
```

```
AAV32615
ID   AAV32615 standard; cDNA; 1001 BP.
XX
AC   AAV32615;
XX
DT   26-OCT-1998  (first entry)
XX
DE   Mutant human apopain (caspase-3) pro-enzyme cDNA.
XX
KW   Mutant human apopain pro-enzyme; caspase-3; protease; SPA;
KW   scintillation proximity binding assay; diabetes; cancer;
KW   4-phosphono(difluoromethyl) phenylalanine group; osteoporosis;
KW   tyrosine phosphatase; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             64..897
FT                   /*tag=  a
FT                   /product= "Mutant apopain C163S pro-enzyme"
FT   misc_feature    148..588
FT                   /*tag=  b
FT                   /note= "Encodes the P17 subunit"
FT   mutation        550..552
FT                   /*tag=  c
FT                   /note= "Changed from TGC in wild-type to TCC in
FT                   mutant"
FT   misc_feature    589..894
FT                   /*tag=  d
FT                   /note= "Encodes the P12 subunit"
XX
PN   WO9820156-A1.
XX
PD   14-MAY-1998.
XX
PF   03-NOV-1997;   97WO-CA00825.
XX
PR   04-NOV-1996;   96US-0030408.
XX
PA   (MERI ) MERCK FROSST CANADA INC.
XX
PI   Desmarais S,  Ford-Hutchinson A,  Friesen R,  Gresser M;
PI   Kennedy B,  Nicholson D,  Ramachandran C,  Skorey K;
XX
DR   WPI; 1998-348103/30.
XX
PT   Enzyme binding assay for detection of useful compounds - uses mutant
PT   form of wild-type enzyme, in which serine replaces cysteine at
PT   active site, to reduce interference from oxidising/alkylating agents
XX
PS   Disclosure; Pages 38-39; 63pp; English.
XX
CC   The present sequence represents a mutated human apopain (caspase-3)
CC   pro-enzyme cDNA.  The cDNA encodes a mutant human apopain C163S
CC   pro-enzyme.  The invention claims for a method for use in a
CC   scintillation proximity binding assay (SPA) for proteases and
CC   phosphatases.  The method involves using mutated proteases and
CC   phosphatases whereby the catalytic cysteine residue of the enzymes
CC   are replaced with a serine or alanine residue to correct the problem of
CC   interference in SPA from extraneous oxidising and alkylating agents.
CC   The mutation affects the catalytic properties of the enzyme but does
CC   not affect their binding properties.  The invention provides new
CC   ligands for use in SPA which have increased binding affinity for a
CC   tyrosine phosphatase or cysteine protease.  The ligands contain at least
CC   two 4-phosphono(difluoromethyl) phenylalanine groups which increase
CC   binding affinity of the ligand to its respective enzyme.  The assay can
CC   be used to determine the ability of new ligands and compound mixtures to
CC   competitively bind with an enzyme.  The method is claimed to allow
CC   a better usage of SPA in the discovery of compounds for the treatment
CC   and study of diseases, e.g. diabetes, cancer and osteoporosis.
XX
SQ   Sequence 1001 BP; 310 A; 172 C; 226 G; 293 T; 0 other;
```

```
  Query Match             69.8%;  Score 809;  DB 19;  Length 1001;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches  909;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49 ttaataaaggtatccatggagaacactgaaaactcagtggattcaaaatccattaaaaat 108

Qy     268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     109 ttggaaccaaagatcatacatggaagcgaatcaatggactctggaatatccctggacaac 168

Qy     328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     169 agttataaaatggattatcctgagatgggtttatgtataataattaataataagaatttt 228

Qy     388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     229 cataagagcactggaatgacatctcggtctggtacagatgtcgatgcagcaaacctcagg 288

Qy     448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     289 gaaacattcagaaacttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaa 348

Qy     508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     349 attgtggaattgatgcgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtt 408

Qy     568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     409 tgtgtgcttctgagccatggtgaagaaggaataatttttggaacaaatggacctgttgac 468

Qy     628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     469 ctgaaaaaaataacaaacttttttcagaggggatcgttgtagaagtctaactggaaaaccc 528

Qy     688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     529 aaacttttcattattcaggcctcccgtggtacagaactggactgtggcattgagacagac 588

Qy     748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     589 agtggtgttgatgatgacatggcgtgtcataaaatacCagtggaggccgacttcttgtat 648

Qy     808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     649 gcatactccacagcacctggttattattcttggcgaaattcaaaggatggctcctggttc 708

Qy     868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     709 atccagtcgctttgtgccatgctgaaacagtatgccgacaagcttgaatttatgcacatt 768

Qy     928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     769 cttacccgggttaaccgaaaggtggcaacagaatttgagtcctttTccTttgacgctact 828

Qy     988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     829 tttcatgcaaagaaacagattccatgtattgtttccatgctcacaaaagaactctatttt 888

Qy    1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     889 tatcactaaagaaatggttggttggtggttttttttagtttgtatgccaagtgagaagat 948

Qy    1108 GGTATATTTGG 1118
           ||||||||||||
Db     949 ggtatatttgg 959


RESULT  14
AAX81219
```

```
ID   AAX81219 standard; DNA; 835 BP.
XX
AC   AAX81219;
XX
DT   10-SEP-1999  (first entry)
XX
DE   Nucleotide sequence of caspase-3 (CPP32).
XX
KW   Rev-caspase; cysteine protease; zymogen; caspase; cancer; gene therapy;
KW   autoimmune disease; caspase-mediated apoptosis; neurodegenerative;
KW   tumour cell; myocardial infarction; human; ds.
XX
OS   Homo sapiens.
XX
PN   WO9935277-A2.
XX
PD   15-JUL-1999.
XX
PF   11-JAN-1999;   99WO-US00632.
XX
PR   09-JAN-1998;   98US-0070987.
XX
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES;
XX
DR   WPI; 1999-419353/35.
DR   P-PSDB; AAY21717.
XX
PT   New isolated nucleic acid molecule encoding a rev-caspase  - used
PT   for screening and identifying inhibitors or enhancers for treating
PT   cancer or autoimmune disease
XX
PS   Disclosure; Fig 13A-B; 74pp; English.
XX
CC   The invention relates to nucleic acid molecules encoding rev-caspases.
CC   Rev-caspases are cysteine proteases that specifically cleave proteins
CC   after Asp residues and is expressed as a zymogen, in which the small
CC   subunit is N-terminal to a large subunit. A gene delivery vehicle
CC   comprising a rev-caspase coding sequence is useful for the treatment of
CC   cancer, where the gene delivery vehicle is internalised by tumour cells.
CC   The gene delivery vehicle can also be used to treat autoimmune diseases.
CC   Cells transfected with a rev-caspase expressing vector can be used in
CC   identification of inhibitors or enhancers of caspase-mediated apoptosis.
CC   In vitro translated rev-caspase can be used to identify an inhibitor or
CC   enhancer of caspase processing activity. Caspase inhibitors are useful
CC   for treating neurodegenerative diseases as well as for inhibiting
CC   apoptosis in the heart following myocardial infarction. Sequences
CC   AAX81217 -AAX81226 represent human caspase genes encoding caspase 1-10
CC   gene products (AAY21715-Y21724).
XX
SQ   Sequence 835 BP; 274 A; 147 C; 178 G; 235 T; 1 other;
```

```
  Query Match             67.6%;  Score 783;  DB 20;  Length 835;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  833;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     223 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATC 282
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 atggagaacactgaaaactcagtggattcaaaatccattaaaaatttggaaccaaagatc 60

Qy     283 ATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGAT 342
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 atacatggaagcgaatcaatggactctggaatatccctggacaacagttataaaatggat 120

Qy     343 TATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGA 402
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 tatcctgagatgggtttatgtataataattaataataagaattttcataaaagcactgga 180

Qy     403 ATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAAC 462
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 atgacatctcggtctggtacagatgtcgatgcagcaaacctcagggaaacattcagaaac 240
```

```
Qy     463 TTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 522
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaaattgtggaattgatg 300

Qy     523 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGC 582
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 cgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtttgtgtgcttctgagc 360

Qy     583 CATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACA 642
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 catggtgaagaaggaataatttttggaacaaatggacctgttgacctgaaaaaaataaca 420

Qy     643 AACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATT 702
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 aacttttcagaggggatcgttgtagaagtctaactggaaaacccaaacttttcattatt 480

Qy     703 CAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGAT 762
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 caggcctgccgtggtacagaactggactgtggcattgagacagacagtggtgttgatgat 540

Qy     763 GACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 822
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     541 gacatggcgtgtcataaaatacca gtggatgccgacttcttgtatgcatactccacagca 600

Qy     823 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGT 882
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 cctggttattattcttggcgaaattcaaaggatggctcctggttcatccagtcgctttgt 660

Qy     883 GCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAAC 942
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 gccatgctgaaacagtatgccgacaagcttgaatttatgcacattcttacccgggttaac 720

Qy     943 CGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAA 1002
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 cgaaaggtggcaacagaatttgagtcctttt ccttt gacgctacttttcatgcaaagaaa 780

Qy    1003 CAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAA 1056
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 cagattccatgtattgtttccatgctcacaaaagaactctatttttatcactaa 834
```

RESULT 15
AAD13084
ID AAD13084 standard; DNA; 7247 BP.
XX
AC AAD13084;
XX
DT 16-OCT-2001 (first entry)
XX
DE pcDNA3-Ub-Met-C3 construct.
XX
KW Protein destabilisation; linker moiety; reporter moiety; disease model;
KW linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW transgenic animal; transgenic plant; disease resistance; caspase-3;
KW human; ubiquitin; cyclic; circular; plasmid pcDNA3-Ub-Met-C3;
KW fusion protein; ds.
XX
OS Chimeric - Mammalia.
OS Chimeric - Homo sapiens.
XX
PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US03791.
XX
PR 04-FEB-2000; 2000US-0498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by
PT -NH-ubiquitin protein endoproteases, comprising two copies of the
PT domain -
XX
PS Example 4; Page 140-141; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method
CC is also useful for developing novel assays for a wide range of
CC post-translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a pcDNA3-Ub-Met-C3 construct comprising wildtype multimerised
CC human ubiquitin destabilisation domain-Mammalia caspase-3 fusion protein.
XX
SQ Sequence 7247 BP; 1794 A; 1796 C; 1826 G; 1831 T; 0 other;

Query Match 65.0%; Score 753; DB 22; Length 7247;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 803; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy     250 TCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCT 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1171 tcaaaatccattaaaaatttggaaccaaagatcatacatggaagcgaatcaatggactct 1230

Qy     310 GGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1231 ggaatatccctggacaacagttataaaatggattatcctgagatgggtttatgtataata 1290

Qy     370 ATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTC 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1291 attaataataagaattttcataaaagcactggaatgacatctcggtctggtacagatgtc 1350

Qy     430 GATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAAT 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1351 gatgcagcaaacctcagggaaacattcagaaacttgaaatatgaagtcaggaataaaaat 1410

Qy     490 GATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1411 gatcttacacgtgaagaaattgtggaattgatgcgtgatgtttctaaagaagatcacagc 1470

Qy     550 AAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGA 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1471 aaaaggagcagttttgtttgtgtgcttctgagccatggtgaagaaggaataatttttgga 1530

Qy     610 ACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGA 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1531 acaaatggacctgttgacctgaaaaaaataacaaacttttcagaggggatcgttgtaga 1590

Qy     670 AGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGAC 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1591 agtctaactggaaaacccaaacttttcattattcaggcctgccgtggtacagaactggac 1650

Qy     730 TGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTG 789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1651 tgtggcattgagacagacagtggtgttgatgatgacatggcgtgtcataaaatacca gtg 1710

Qy     790 GAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCA 849
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1711 gatgccgacttcttgtatgcatactccacagcacctggttattattcttggcgaaattca 1770
```

```
Qy     850 AAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAG 909
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1771 aaggatggctcctggttcatccagtcgctttgtgccatgctgaaacagtatgccgacaag 1830

Qy     910 CTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCC 969
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1831 cttgaatttatgcacattcttacccgggttaaccgaaaggtggcaacagaatttgagtcc 1890

Qy     970 TTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTC 1029
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1891 ttttccttgacgctacttttcatgcaaagaaacagattccatgtattgtttccatgctc 1950

Qy    1030 ACAAAAGAACTCTATTTTTATCAC 1053
           ||||||||||||||||||||||||
Db    1951 acaaaagaactctatttttatcac 1974
```

Search completed: August 16, 2002, 01:49:05
Job time: 6206 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 01:43:40 ; Search time 3312.14 Seconds
(without alignments)
7322.715 Million cell updates/sec

Title: US-09-613-508-3
Perfect score: 1159
Sequence: 1 GCACGAGCGGATGGGTGCTA.....GGGGACCTACTCTCATGCTG 1159

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1159	100.0	1159	6	AR103059	AR103059 Sequence
2	1031	89.0	2652	9	BC016926	BC016926 Homo sapi
3	1004	86.6	2635	6	AR172878	AR172878 Sequence
4	1004	86.6	2635	6	AX134409	AX134409 Sequence
5	1004	86.6	2635	6	AX331135	AX331135 Sequence
6	1004	86.6	2635	9	HSU13737	U13737 Human cyste
7	860	74.2	1001	6	AR109589	AR109589 Sequence
8	860	74.2	1001	9	HSU13738	U13738 Human cyste
9	834	72.0	834	6	AX207737	AX207737 Sequence
10	834	72.0	834	9	HSU26943	U26943 Human cyste
11	825	71.2	6436	6	AX207740	AX207740 Sequence
12	825	71.2	6703	6	AX207741	AX207741 Sequence
13	825	71.2	6946	6	AX207742	AX207742 Sequence
14	825	71.2	7189	6	AX207743	AX207743 Sequence
15	809	69.8	1001	6	AR109591	AR109591 Sequence
16	753	65.0	7248	6	AX207746	AX207746 Sequence
17	677	58.4	834	9	HSA413269	AJ413269 Homo sapi
18	501	43.2	1673	6	AX260262	AX260262 Sequence
19	501	43.2	2566	6	AX260259	AX260259 Sequence
20	476	41.1	873	12	AF052645	AF052645 Synthetic
21	438	37.8	3614	6	AX235203	AX235203 Sequence
22	397	34.3	858	12	AF052647	AF052647 Synthetic
23	387	33.4	786	6	AX134427	AX134427 Sequence
c 24	293	25.3	169950	2	AC084049	AC084049 Homo sapi
25	293	25.3	207425	2	AC060790	AC060790 Homo sapi
c 26	160	13.8	178290	2	AC079257	AC079257 Homo sapi
27	52	4.5	834	4	AB029345	AB029345 Sus scrof
c 28	50	4.3	131489	9	AL592104	AL592104 Human DNA
c 29	50	4.3	197371	2	AL355524	AL355524 Homo sapi
30	44	3.8	53	6	AX207744	AX207744 Sequence
31	43	3.7	50	6	AX207738	AX207738 Sequence
c 32	41	3.5	48	6	AX207739	AX207739 Sequence
c 33	41	3.5	48	6	AX207745	AX207745 Sequence
34	38	3.3	665	4	AF068837	AF068837 Ovis arie
35	35	3.0	400	11	G13613	G13613 human STS S
36	34	2.9	144682	2	AC094580	AC094580 Rattus no
37	34	2.9	170674	2	AC096596	AC096596 Rattus no
38	33	2.8	2395	10	CGU27463	U27463 Cricetinae
39	32	2.8	53	6	AR053334	AR053334 Sequence
40	32	2.8	53	6	AR053338	AR053338 Sequence
c 41	31	2.7	51	6	AR053337	AR053337 Sequence
c 42	31	2.7	51	6	AR053341	AR053341 Sequence
43	30	2.6	90	6	AR146928	AR146928 Sequence
44	30	2.6	90	6	AR146940	AR146940 Sequence
c 45	29	2.5	52	6	AR053335	AR053335 Sequence

ALIGNMENTS

RESULT 1
AR103059
LOCUS AR103059 1159 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 3 from patent US 6087150.
ACCESSION AR103059
VERSION AR103059.1 GI:12814647
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1159)
AUTHORS He,W.W., Rosen,C.A., Hudson,P.L. and Hastings,G.A.
TITLE Interleukin-1 .beta. converting enzyme like apoptosis protease-3 and 4
JOURNAL Patent: US 6087150-A 3 11-JUL-2000;
FEATURES Location/Qualifiers
source 1. .1159
/organism="unknown"
BASE COUNT 329 a 220 c 276 g 334 t
ORIGIN

Query Match 100.0%; Score 1159; DB 6; Length 1159;

```
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1159;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60

Qy      61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120

Qy     121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180

Qy     181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240

Qy     241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300

Qy     301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360

Qy     361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420

Qy     421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480

Qy     481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540

Qy     541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600

Qy     601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660

Qy     661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720

Qy     721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780

Qy     781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840

Qy     841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900

Qy     901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960

Qy     961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020

Qy    1021 TCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 TCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080

Qy    1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTG 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTG 1140

Qy    1141 GGGACCTACTCTCATGCTG 1159
           ||||||||||||||||||||
Db    1141 GGGACCTACTCTCATGCTG 1159


RESULT   2
BC016926
LOCUS       BC016926               2652 bp    mRNA    linear   PRI 09-NOV-2001
DEFINITION  Homo sapiens, clone MGC:21241 IMAGE:4419175, mRNA, complete cds.
ACCESSION   BC016926
VERSION     BC016926.1  GI:16877343
KEYWORDS    MGC.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2652)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 27 Row: j Column: 23
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 14790118.
FEATURES             Location/Qualifiers
     source          1. .2652
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="MGC:21241 IMAGE:4419175"
                     /tissue_type="Lymph, lymphoma"
                     /clone_lib="NIH_MGC_85"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             228. .1061
                     /codon_start=1
                     /product="Unknown (protein for MGC:21241)"
                     /protein_id="AAH16926.1"
                     /db_xref="GI:16877344"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVEADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
BASE COUNT      824 a    466 c    576 g    786 t
ORIGIN


  Query Match             89.0%;  Score 1031;  DB 9;  Length 2652;
  Best Local Similarity   99.9%;  Pred. No. 0;
```

```
  Matches 1081;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      37 GAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTG 96
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42 GAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTG 101

Qy      97 GTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGC 156
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102 GTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGC 161

Qy     157 GCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAG 216
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     162 GCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCTCTTCCCCCATTCTCATTAATAAAG 221

Qy     217 GTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCA 276
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 GTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCA 281

Qy     277 AAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAA 336
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     282 AAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAA 341

Qy     337 ATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGC 396
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     342 ATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGC 401

Qy     397 ACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTC 456
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     402 ACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTC 461

Qy     457 AGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAA 516
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     462 AGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAA 521

Qy     517 TTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTT 576
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     522 TTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTT 581

Qy     577 CTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAA 636
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     582 CTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAA 641

Qy     637 ATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTC 696
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     642 ATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTC 701

Qy     697 ATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTT 756
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     702 ATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTT 761

Qy     757 GATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCC 816
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     762 GATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCC 821

Qy     817 ACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCG 876
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     822 ACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCG 881

Qy     877 CTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGG 936
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     882 CTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGG 941

Qy     937 GTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCA 996
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     942 GTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCA 1001

Qy     997 AAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAA 1056
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1002 AAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAA 1061

Qy    1057 AGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTT 1116
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1062 AGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTT 1121

Qy    1117 GG 1118
           ||
Db    1122 GG 1123


RESULT   3
AR172878
LOCUS       AR172878               2635 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 3 from patent US 6303374.
ACCESSION   AR172878
VERSION     AR172878.1  GI:17912369
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2635)
  AUTHORS   Zhang,H. and Cowsert,L.M.
  TITLE     Antisense modulation of caspase 3 expression
  JOURNAL   Patent: US 6303374-A 3 16-OCT-2001;
FEATURES             Location/Qualifiers
     source          1. .2635
                     /organism="unknown"
BASE COUNT      815 a    465 c    569 g    786 t
ORIGIN


  Query Match             86.6%;  Score 1004;  DB 6;  Length 2635;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1104;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      13 GGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTG 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 GGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTG 74

Qy      73 GCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGT 132
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 GCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGT 134

Qy     133 TAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 TAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCT 194

Qy     193 CTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCA 252
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     195 CTTCCCCCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCA 254

Qy     253 AAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGA 312
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     255 AAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGA 314

Qy     313 ATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     315 ATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

Qy     373 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGAT 432
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGAT 434

Qy     433 GCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGAT 492
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     435 GCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGAT 494

Qy     493 CTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAA 552
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     495 CTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAA 554

Qy     553 AGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACA 612
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     555 AGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACA 614

Qy     613 AATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 672
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    615 AATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 674

Qy    673 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGT 732
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGT 734

Qy    733 GGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAG 792
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    735 GGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAT 794

Qy    793 GCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAG 852
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    795 GCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAG 854

Qy    853 GATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTT 912
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    855 GATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTT 914

Qy    913 GAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 972
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    915 GAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 974

Qy    973 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACA 1032
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    975 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACA 1034

Qy   1033 AAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTAT 1092
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1035 AAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTAT 1094

Qy   1093 GCCAAGTGAGAAGATGGTATATTTGG 1118
          |||||||||||||||||||||||||
Db   1095 GCCAAGTGAGAAGATGGTATATTTGG 1120


RESULT   4
AX134409
LOCUS       AX134409                2635 bp    DNA     linear   PAT 29-MAY-2001
DEFINITION  Sequence 5 from Patent WO0129232.
ACCESSION   AX134409
VERSION     AX134409.1  GI:14271055
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2635)
  AUTHORS   Cordell,B. and Li,Y.
  TITLE     Functional cloning of genes encoding proteins/enzymes involved in
            proteolytic cleavage
  JOURNAL   Patent: WO 0129232-A 5 26-APR-2001;
            Scios Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2635
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             225. .1058
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC39521.1"
                     /db_xref="GI:14271056"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVDADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
BASE COUNT      815 a    465 c    569 g    786 t
ORIGIN


  Query Match           86.6%;  Score 1004;  DB 6;  Length 2635;
  Best Local Similarity 99.8%;  Pred. No. 0;

  Matches 1104;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     13 GGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTG 72
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 GGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTG 74

Qy     73 GCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGT 132
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 GCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGT 134

Qy    133 TAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCT 192
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 TAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCT 194

Qy    193 CTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCA 252
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 CTTCCCCCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCA 254

Qy    253 AAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGA 312
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 AAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGA 314

Qy    313 ATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 372
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 ATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

Qy    373 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGAT 432
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGAT 434

Qy    433 GCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGAT 492
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    435 GCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGAT 494

Qy    493 CTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAA 552
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    495 CTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAA 554

Qy    553 AGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACA 612
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    555 AGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACA 614

Qy    613 AATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 672
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    615 AATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 674

Qy    673 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGT 732
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGT 734

Qy    733 GGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAG 792
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    735 GGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAT 794

Qy    793 GCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAG 852
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    795 GCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAG 854

Qy    853 GATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTT 912
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    855 GATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTT 914

Qy    913 GAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 972
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    915 GAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 974

Qy    973 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACA 1032
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    975 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACA 1034

Qy   1033 AAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTAT 1092
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1035 AAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTAT 1094
```

```
Qy   1093 GCCAAGTGAGAAGATGGTATATTTGG 1118
          ||||||||||||||||||||||||||
Db   1095 GCCAAGTGAGAAGATGGTATATTTGG 1120


RESULT   5
AX331135
LOCUS       AX331135                2635 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 1644 from Patent WO0194629.
ACCESSION   AX331135
VERSION     AX331135.1  GI:18121769
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 1644 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .2635
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      815 a    465 c    569 g    786 t
ORIGIN


  Query Match             86.6%;  Score 1004;  DB 6;  Length 2635;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1104;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     13 GGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTG 72
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 GGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTG 74

Qy     73 GCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGT 132
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 GCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGT 134

Qy    133 TAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCT 192
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 TAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCT 194

Qy    193 CTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCA 252
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 CTTCCCCCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCA 254

Qy    253 AAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGA 312
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 AAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGA 314

Qy    313 ATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 372
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 ATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

Qy    373 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGAT 432
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGAT 434

Qy    433 GCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGAT 492
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    435 GCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGAT 494

Qy    493 CTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAA 552
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    495 CTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAA 554

Qy    553 AGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACA 612
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    555 AGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACA 614

Qy    613 AATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 672
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    615 AATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 674

Qy    673 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGT 732
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGT 734

Qy    733 GGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAG 792
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    735 GGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAT 794

Qy    793 GCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAG 852
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    795 GCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAG 854

Qy    853 GATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTT 912
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    855 GATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTT 914

Qy    913 GAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 972
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    915 GAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 974

Qy    973 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACA 1032
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    975 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACA 1034

Qy   1033 AAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTAT 1092
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1035 AAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTAT 1094

Qy   1093 GCCAAGTGAGAAGATGGTATATTTGG 1118
          ||||||||||||||||||||||||||
Db   1095 GCCAAGTGAGAAGATGGTATATTTGG 1120


RESULT   6
HSU13737
LOCUS       HSU13737                2635 bp    mRNA    linear   PRI 14-APR-1995
DEFINITION  Human cysteine protease CPP32 isoform alpha mRNA, complete cds.
ACCESSION   U13737
VERSION     U13737.1  GI:561665
KEYWORDS    cysteine protease; interleukin 1-beta converting enzyme; apoptotic
            protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2635)
  AUTHORS   Fernandes-Alnemri,T., Litwack,G. and Alnemri,E.S.
  TITLE     CPP32, a novel human apoptotic protein with homology to
            Caenorhabditis elegans cell death protein Ced-3 and mammalian
            interleukin-1 beta-converting enzyme
  JOURNAL   J. Biol. Chem. 269 (49), 30761-30764 (1994)
  MEDLINE   95074098
REFERENCE   2  (bases 1 to 2635)
  AUTHORS   Fernandes-Alnemri,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-AUG-1994) Teresa Fernandes-Alnemri, Pharmacology,
            Thomas Jefferson University, Jefferson Cancer Institute, 233, S.
            Tenth Street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .2635
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="J1.1 EcoRI/XhoI fragment"
                     /cell_line="Jurkat"
                     /cell_type="T-lymphocyte"
                     /clone_lib="Human Jurkat T-lymphocyte lambda Uni-ZAP XR
```

```
                cDNA library"
     CDS        225. .1058
                /note="new member of the interleukin 1-beta converting
                enzyme gene family of cysteine proteases"
                /codon_start=1
                /product="cysteine protease CPP32 isoform alpha"
                /protein_id="AAA65015.1"
                /db_xref="GI:561666"
                /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                QACRGTELDCGIETDSGVDDDMACHKIPVDADFLYAYSTAPGYYSWRNSKDGSWFIQS
                LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                H"
     polyA_site 2635
                /note="19 A residues"
BASE COUNT      815 a    465 c    569 g    786 t
ORIGIN
```

Query Match 86.6%; Score 1004; DB 9; Length 2635;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1104; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      13 GGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTG 72
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 GGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTG 74

Qy      73 GCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGT 132
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 GCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGGAGT 134

Qy     133 TAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCT 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 TAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCT 194

Qy     193 CTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCA 252
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     195 CTTCCCCCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCA 254

Qy     253 AAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGA 312
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     255 AAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGA 314

Qy     313 ATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     315 ATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

Qy     373 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGAT 432
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGAT 434

Qy     433 GCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGAT 492
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     435 GCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGAT 494

Qy     493 CTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAA 552
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     495 CTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAA 554

Qy     553 AGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACA 612
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     555 AGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACA 614

Qy     613 AATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 672
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     615 AATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 674

Qy     673 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGT 732
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGT 734

Qy     733 GGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAG 792
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     735 GGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAT 794

Qy     793 GCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAG 852
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     795 GCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAG 854

Qy     853 GATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTT 912
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     855 GATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTT 914

Qy     913 GAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 972
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     915 GAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 974

Qy     973 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACA 1032
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     975 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACA 1034

Qy    1033 AAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTAT 1092
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1035 AAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTAT 1094

Qy    1093 GCCAAGTGAGAAGATGGTATATTTGG 1118
           |||||||||||||||||||||||||
Db    1095 GCCAAGTGAGAAGATGGTATATTTGG 1120
```

```
RESULT   7
AR109589
LOCUS       AR109589                1001 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 5 from patent US 6114132.
ACCESSION   AR109589
VERSION     AR109589.1  GI:12825865
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1001)
  AUTHORS   Desmarais,S., Gresser,M., Friesen,R., Kennedy,B., Nicholson,D.,
            Skorey,K., Ramachandran,C. and Ford-Hutchinson,A.
  TITLE     Prosphatase binding assay
  JOURNAL   Patent: US 6114132-A 5 05-SEP-2000;
FEATURES             Location/Qualifiers
     source          1. .1001
                     /organism="unknown"
BASE COUNT      310 a    171 c    227 g    293 t
ORIGIN
```

Query Match 74.2%; Score 860; DB 6; Length 1001;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy     208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy     268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy     328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy     388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy     448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348
```

```
Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    529 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy    988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy   1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy   1108 GGTATATTTGG 1118
          ||||||||||||
Db    949 GGTATATTTGG 959
```

```
RESULT   8
HSU13738
LOCUS       HSU13738                1001 bp    mRNA    linear   PRI 14-APR-1995
DEFINITION  Human cysteine protease CPP32 isoform beta mRNA, complete cds.
ACCESSION   U13738
VERSION     U13738.1  GI:561667
KEYWORDS    cysteine protease; interleukin 1-beta converting enzyme; apoptotic
            protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1001)
  AUTHORS   Fernandes-Alnemri,T., Litwack,G. and Alnemri,E.S.
  TITLE     CPP32, a novel human apoptotic protein with homology to
            Caenorhabditis elegans cell death protein Ced-3 and mammalian
            interleukin-1 beta-converting enzyme
  JOURNAL   J. Biol. Chem. 269 (49), 30761-30764 (1994)
  MEDLINE   95074098
REFERENCE   2  (bases 1 to 1001)
  AUTHORS   Fernandes-Alnemri,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-AUG-1994) Teresa Fernandes-Alnemri, Pharmacology,
            Thomas Jefferson University, Jefferson Cancer Institute, 233, S.
            Tenth Street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .1001
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="J6.1 Pst I fragment"
                     /cell_line="Jurkat"
                     /cell_type="T-lymphocyte"
                     /clone_lib="Human Jurkat T-lymphocyte lambda Uni-ZAP XR
                     cDNA library"
     5'UTR           1. .63
                     /note="An alternatively spliced isoform which lacks most
                     of the 5' nontranslated sequence"
     CDS             64. .897
                     /note="new member of the interleukin 1-beta converting
                     enzyme gene family of cysteine proteases"
                     /codon_start=1
                     /product="cysteine protease CPP32 isoform beta"
                     /protein_id="AAB60355.1"
                     /db_xref="GI:561668"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVEADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
     misc_difference 234
                     /note="differs from isoform alpha by one bp substitution
                     at position 234"
                     /replace="a"
     misc_difference 633
                     /note="differs from isoform alpha by one bp substitution
                     at position 633"
                     /replace="t"
BASE COUNT      310 a    171 c    227 g    293 t
ORIGIN
```

Query Match 74.2%; Score 860; DB 9; Length 1001;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    529 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy    988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy   1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy   1108 GGTATATTTGG 1118
          ||||||||||||
Db    949 GGTATATTTGG 959


RESULT  9
AX207737
LOCUS       AX207737                834 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 35 from Patent WO0157242.
ACCESSION   AX207737
VERSION     AX207737.1  GI:15422421
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 834)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 35 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .834
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      274 a    147 c    179 g    234 t
ORIGIN


  Query Match             72.0%;  Score 834;  DB 6;  Length 834;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  834;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    223 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATC 282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATC 60

Qy    283 ATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGAT 342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGAT 120

Qy    343 TATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGA 402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGA 180

Qy    403 ATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAAC 462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    181 ATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAAC 240

Qy    463 TTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 300

Qy    523 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGC 582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGC 360

Qy    583 CATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACA 642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACA 420

Qy    643 AACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATT 702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATT 480

Qy    703 CAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGAT 762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGAT 540

Qy    763 GACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 600

Qy    823 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGT 882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGT 660

Qy    883 GCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAAC 942
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAAC 720

Qy    943 CGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAA 1002
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAA 780

Qy   1003 CAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAA 1056
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAA 834


RESULT  10
HSU26943
LOCUS       HSU26943                834 bp    mRNA    linear   PRI 29-AUG-1995
DEFINITION  Human cysteine protease Yama mRNA, complete cds.
ACCESSION   U26943
VERSION     U26943.1  GI:857568
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 834)
  AUTHORS   Tewari,M., Quan,L.T., O'Rourke,K., Desnoyers,S., Zeng,Z.,
            Beidler,D.R., Poirier,G.G., Salvesen,G.S. and Dixit,V.M.
  TITLE     Yama/CPP32 beta, a mammalian homolog of CED-3, is a
            CrmA-inhibitable protease that cleaves the death substrate
            poly(ADP-ribose) polymerase
  JOURNAL   Cell 81 (5), 801-809 (1995)
  MEDLINE   95292347
REFERENCE   2  (bases 1 to 834)
  AUTHORS   Tewari,M., O'Rourke,K., Lin,Z. and Dixit,V.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-1995) Muneesh Tewari, Pathology, University of
            Michigan, 1301 Catherine, Ann Arbor, MI 48109, USA
FEATURES             Location/Qualifiers
     source          1. .834
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="umbilical vein endothelial cells treated with
                     tumor necrosis factor and cycloheximide"
```

```
     CDS             1. .834
                     /note="cysteine protease CPP32beta; CED-3/ICE protein
                     family"
                     /codon_start=1
                     /product="Yama protein"
                     /protein_id="AAA74929.1"
                     /db_xref="GI:857569"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVEADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
BASE COUNT      274 a    147 c    179 g    234 t
ORIGIN

  Query Match             72.0%;  Score 834;  DB 9;  Length 834;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  834;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    223 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATC 282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATC 60

Qy    283 ATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGAT 342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGAT 120

Qy    343 TATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGA 402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGA 180

Qy    403 ATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAAC 462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAAC 240

Qy    463 TTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 300

Qy    523 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGC 582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGC 360

Qy    583 CATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACA 642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACA 420

Qy    643 AACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATT 702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATT 480

Qy    703 CAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGAT 762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGAT 540

Qy    763 GACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 600

Qy    823 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGT 882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGT 660

Qy    883 GCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAAC 942
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAAC 720

Qy    943 CGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAA 1002
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAA 780

Qy   1003 CAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAA 1056
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAA 834


RESULT  11
AX207740
LOCUS       AX207740               6436 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 38 from Patent WO0157242.
ACCESSION   AX207740
VERSION     AX207740.1  GI:15422424
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 6436)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 38 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6436
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-1XUb-C3 construct"
BASE COUNT     1577 a   1606 c   1607 g   1646 t
ORIGIN

  Query Match             71.2%;  Score 825;  DB 6;  Length 6436;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  825;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    229 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 288
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1144 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 1203

Qy    289 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 348
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1204 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 1263

Qy    349 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 408
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1264 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 1323

Qy    409 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 468
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1324 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 1383

Qy    469 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 528
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1384 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1443

Qy    529 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 588
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1444 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 1503

Qy    589 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 648
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1504 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 1563

Qy    649 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 708
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1564 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 1623

Qy    709 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 768
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1624 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 1683

Qy    769 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 828
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1684 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 1743
```

```
Qy    829 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 888
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1744 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 1803

Qy    889 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1804 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 1863

Qy    949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1008
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1864 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1923

Qy   1009 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1053
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1924 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1968


RESULT  12
AX207741
LOCUS       AX207741                6703 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 39 from Patent WO0157242.
ACCESSION   AX207741
VERSION     AX207741.1  GI:15422425
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 6703)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 39 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6703
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-2XUb-C3 construct"
BASE COUNT     1653 a   1677 c   1674 g   1699 t
ORIGIN


  Query Match           71.2%;  Score 825;  DB 6;  Length 6703;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  825;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    229 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 288
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1411 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 1470

Qy    289 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 348
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1471 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 1530

Qy    349 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 408
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1531 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 1590

Qy    409 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 468
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1591 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 1650

Qy    469 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 528
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1651 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1710

Qy    529 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 588
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1711 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 1770

Qy    589 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 648
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1771 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 1830

Qy    649 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 708
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1831 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 1890

Qy    709 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 768
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1891 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 1950

Qy    769 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 828
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1951 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2010

Qy    829 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 888
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2011 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 2070

Qy    889 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2071 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 2130

Qy    949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1008
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2131 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 2190

Qy   1009 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1053
          |||||||||||||||||||||||||||||||||||||||||||||
Db   2191 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 2235


RESULT  13
AX207742
LOCUS       AX207742                6946 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 40 from Patent WO0157242.
ACCESSION   AX207742
VERSION     AX207742.1  GI:15422426
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 6946)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 40 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6946
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-3XUb-C3 construct"
BASE COUNT     1723 a   1742 c   1735 g   1746 t
ORIGIN


  Query Match           71.2%;  Score 825;  DB 6;  Length 6946;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  825;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    229 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 288
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1654 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 1713

Qy    289 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 348
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1714 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 1773

Qy    349 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 408
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1774 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 1833

Qy    409 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 468
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1834 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 1893
```

```
Qy    469 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 528
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1894 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1953

Qy    529 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 588
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1954 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 2013

Qy    589 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 648
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2014 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 2073

Qy    649 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 708
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2074 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 2133

Qy    709 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 768
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2134 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 2193

Qy    769 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 828
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2194 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2253

Qy    829 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 888
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2254 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 2313

Qy    889 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2314 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 2373

Qy    949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1008
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2374 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 2433

Qy   1009 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1053
          |||||||||||||||||||||||||||||||||||||||||||||
Db   2434 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 2478
```

RESULT 14
AX207743

```
LOCUS       AX207743               7189 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 41 from Patent WO0157242.
ACCESSION   AX207743
VERSION     AX207743.1  GI:15422427
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 7189)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 41 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7189
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-4XUb-C3 construct"
BASE COUNT     1793 a   1807 c   1796 g   1793 t
ORIGIN
```

Query Match 71.2%; Score 825; DB 6; Length 7189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 825; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    229 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 288
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 1956

Qy    289 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 348
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 2016

Qy    349 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 408
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2017 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 2076

Qy    409 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 468
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2077 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 2136

Qy    469 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 528
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2137 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 2196

Qy    529 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 588
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2197 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 2256

Qy    589 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 648
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2257 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 2316

Qy    649 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 708
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2317 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 2376

Qy    709 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 768
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2377 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 2436

Qy    769 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 828
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2437 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2496

Qy    829 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 888
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2497 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 2556

Qy    889 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2557 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 2616

Qy    949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1008
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2617 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 2676

Qy   1009 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1053
          |||||||||||||||||||||||||||||||||||||||||||||
Db   2677 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 2721
```

RESULT 15
AR109591

```
LOCUS       AR109591               1001 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 9 from patent US 6114132.
ACCESSION   AR109591
VERSION     AR109591.1  GI:12825867
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1001)
  AUTHORS   Desmarais,S., Gresser,M., Friesen,R., Kennedy,B., Nicholson,D.,
            Skorey,K., Ramachandran,C. and Ford-Hutchinson,A.
  TITLE     Prosphatase binding assay
  JOURNAL   Patent: US 6114132-A 9 05-SEP-2000;
FEATURES             Location/Qualifiers
     source          1. .1001
                     /organism="unknown"
BASE COUNT      310 a    172 c    226 g    293 t
ORIGIN
```

Query Match 69.8%; Score 809; DB 6; Length 1001;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 909; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy    628 CTGAAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    529 AAACTTTTCATTATTCAGGCCTCCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy    988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy   1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy   1108 GGTATATTTGG 1118
          ||||||||||| 
Db    949 GGTATATTTGG 959
```

Search completed: August 16, 2002, 01:44:05

Job time: 6331 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 00:05:10 ; Search time 323.26 Seconds
(without alignments)
6155.738 Million cell updates/sec

Title: US-09-613-508-3
Perfect score: 1159
Sequence: 1 GCACGAGCGGATGGGTGCTA.....GGGGACCTACTCTCATGCTG 1159

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1159	100.0	1159	17	AAT15277	Human interleukin-
2	1133.4	97.8	1158	17	AAT33567	Pro-Yama cDNA. Ho
3	1133.4	97.8	1158	19	AAV05471	Nucleic acid encod
4	1118	96.5	2635	22	AAS10503	Human caspase 3 DN
5	1118	96.5	2635	22	AAD03910	Human caspase-3 cD
6	921.6	79.5	1001	19	AAV32608	Mutant human apopa
7	921.6	79.5	1001	19	AAV32615	Mutant human apopa
8	839.6	72.4	2538	21	AAC64043	pYNGCas3p10p35 ins
9	834	72.0	834	22	AAD13075	Mammalia caspase-3
10	832.4	71.8	835	20	AAX81219	Nucleotide sequenc
11	826.4	71.3	7247	22	AAD13084	pcDNA3-Ub-Met-C3 c
12	825	71.2	6436	22	AAD13078	pcDNA3-1XUb-C3 con
13	825	71.2	6703	22	AAD13079	pcDNA3-2XUb-C3 con
14	825	71.2	6946	22	AAD13080	pcDNA3-3XUb-C3 con
15	825	71.2	7189	22	AAD13081	pcDNA3-4XUb-C3 con
16	687.4	59.3	834	22	AAI66510	Pig caspase coding
17	684.6	59.1	786	22	AAD03918	Caspase chimera ca
18	642.4	55.4	1177	19	AAV13959	Rat interleukin-1
19	642.4	55.4	1177	19	AAV13960	Rat interleukin-1
20	631.8	54.5	1297	22	AAS10510	Mouse caspase 3 DN
21	525.8	45.4	873	20	AAX81214	Nucleotide sequenc
22	522.6	45.1	858	20	AAX81215	Nucleotide sequenc
23	438.4	37.8	3614	22	AAH26878	Anti-HER2 Fab reve
24	425.4	36.7	443	21	AAC01186	Human secreted pro
25	240	20.7	2309	18	AAT66992	Apototic protease
26	240	20.7	2309	22	AAD03914	Human caspase-7 DN
27	240	20.7	2341	18	AAT66970	Cysteine protease
28	238.8	20.6	913	20	AAX81223	Nucleotide sequenc
29	229	19.8	1371	17	AAT15276	Human interleukin-
30	181	15.6	298	22	AAF90397	Mouse procaspase-3
31	158.8	13.7	2377	18	AAT66993	Cysteine protease
32	131.6	11.4	2658	20	AAV80061	Aspartate-specific
33	131.4	11.3	883	20	AAX81222	Nucleotide sequenc
34	131.2	11.3	1545	18	AAT43537	Apoptotic cysteine
35	131.2	11.3	1545	22	AAD03913	Human caspase-6 DN
36	126.4	10.9	1183	21	AAZ98834	Interleukin 1 conv
37	126.4	10.9	1204	21	AAZ98839	Interleukin 1 conv
38	109.8	9.5	1265	23	ABL03125	Drosophila melanog
39	109.8	9.5	3265	23	ABL03124	Drosophila melanog
40	98.2	8.5	1313	18	AAT43538	Mch2 beta isoform
41	97	8.4	3008	23	ABL09303	Drosophila melanog
42	96.4	8.3	903	20	AAX81216	Nucleotide sequenc
43	59.4	5.1	376	22	AAI82558	Human polynucleoti
44	55.2	4.8	2619	18	AAT61406	MACH isoform alpha
45	55.2	4.8	2887	18	AAT61405	MACH isoform alpha

ALIGNMENTS

RESULT 1
AAT15277
ID AAT15277 standard; cDNA; 1159 BP.
XX
AC AAT15277;
XX
DT 28-OCT-1996 (first entry)
XX
DE Human interleukin-I-converting enzyme-like apoptosis protease-4 cDNA.
XX
KW ICE-LAP-4; interleukin-I-converting enzyme-like apoptosis protease;
KW enzyme; Alzheimer's disease; Parkinson's disease; septic shock;
KW head injury; rheumatoid arthritis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 115..1137
FT /*tag= a
XX
PN WO9613603-A1.
XX
PD 09-MAY-1996.
XX
PF 06-JUN-1995; 95WO-US07235.
XX
PR 01-NOV-1994; 94US-0334251.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Hastings GA, He W, Hudson PL, Rosen CA;
XX

```
DR   WPI; 1996-239509/24.
DR   P-PSDB; AAR95831.
XX
PT   Human ICE-LAP-3 and -4 DNA and protein  - useful in the diagnosis
PT   and treatment of Alzheimer's disease, Parkinson's disease,
PT   rheumatoid arthritis, septic shock and head injury
XX
PS   Claim 1; Page 44; 67pp; English.
XX
CC   This sequence was derived from a cDNA library from human endometrial
CC   tumuor and may be used for the recombinant production of the ICE-LAP-
CC   4 protein e.g. in COS cells.  The encoded protein may be used
CC   therapeutically, e.g. as an antitumor or antiviral agent and to
CC   control embryonic development and tissue homeostasis.  The protein
CC   can also be used to treat immunosuppression disorders, such as AIDS,
CC   by targeting virus infected cells for cell death.  The DNA may find
CC   use in gene therapy applications.
XX
SQ   Sequence 1159 BP; 329 A; 220 C; 276 G; 334 T; 0 other;

  Query Match           100.0%;  Score 1159;  DB 17;  Length 1159;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1159;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gcacgagcggatgggtgctattgtgaggcggttgtagaagagtttcgtgagtgctcgcag 60

Qy     61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ctcatacctgtggctgtgtatccgtggccacagctggttggcgtcgccttgaaatcccag 120

Qy    121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 gccgtgaggagttagcgagccctgctcacactcggcgctctggttttcggtgggtgtgcc 180

Qy    181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ctgcacctgcctcttcccgcattctcattaataaaggtatccatggagaacactgaaaac 240

Qy    241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 tcagtggattcaaaatccattaaaaatttggaaccaaagatcatacatggaagcgaatca 300

Qy    301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 atggactctggaatatccctggacaacagttataaaatggattatcctgagatgggttta 360

Qy    361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 tgtataataattaataataagaattttcataaaagcactggaatgacatctcggtctggt 420

Qy    421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 acagatgtcgatgcagcaaacctcagggaaacattcagaaacttgaaatatgaagtcagg 480

Qy    481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 aataaaaatgatcttacacgtgaagaaattgtggaattgatgcgtgatgtttctaaagaa 540

Qy    541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 gatcacagcaaaaggagcagttttgtttgtgtgcttctgagccatggtgaagaaggaata 600

Qy    601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 atttttggaacaaatggacctgttgacctgaaaaaaataacaaacttttttcagaggggat 660

Qy    661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 cgttgtagaagtctaactggaaaacccaaacttttcattattcaggcctgccgtggtaca 720

Qy    721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 gaactggactgtggcattgagacagacagtggtgttgatgatgacatggcgtgtcataaa 780

Qy    781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 ataccagtggaggccgacttcttgtatgcatactccacagcacctggttattattcttgg 840

Qy    841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 cgaaattcaaaggatggctcctggttcatccagtcgctttgtgccatgctgaaacagtat 900

Qy    901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 gccgacaagcttgaatttatgcacattcttacccgggttaaccgaaaggtggcaacagaa 960

Qy    961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 tttgagtccttttccttttgacgctacttttcatgcaaagaaacagattccatgtattgtt 1020

Qy   1021 TCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 tccatgctcacaaaagaactctatttttatcactaaagaaatggttggttggtggttttt 1080

Qy   1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 tttagtttgtatgccaagtgagaagatggtatatttgggtactgtatttccctctcattg 1140

Qy   1141 GGGACCTACTCTCATGCTG 1159
          ||||||||||||||||||
Db   1141 gggacctactctcatgctg 1159


RESULT   2
AAT33567
ID   AAT33567 standard; cDNA; 1158 BP.
XX
AC   AAT33567;
XX
DT   06-DEC-1996  (first entry)
XX
DE   Pro-Yama cDNA.
XX
KW   Yama; CrmA; apoptosis; Fas receptor; T-lymphocyte; AIDS; HIV;
KW   human immunodeficiency virus; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             224..1057
FT                   /*tag=  a
XX
PN   WO9625501-A2.
XX
PD   22-AUG-1996.
XX
PF   09-FEB-1996;   96WO-US01882.
XX
PR   01-JUN-1995;   95US-0457731.
PR   13-FEB-1995;   95US-0389812.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Dixit VM;
XX
DR   WPI; 1996-393402/39.
DR   P-PSDB; AAW00677.
XX
PT   New nucleic acid encoding Yama protein or CrmA mutant - useful for
PT   modulating apoptosis, maintaining T cell viability in AIDS patients
PT   and for drug screening
XX
PS   Disclosure; Fig 1; 118pp; English.
```

```
XX
CC   A cDNA clone (AAT33567) codes for pro-Yama (AAW00677), a zymogen which
CC   upon activation cleaves PARP to an 85 kDa form.  Activated Yama has
CC   the ability to modulate cellular function associated with the Fas
CC   receptor pathway such as Fas-associated apoptosis.  This activity is
CC   inhibited by CrmA (see also AAW00678).  The cDNA was isolated from a
CC   cDNA library prepd. from human umbilical vein endothelial cells
CC   treated with tumour necrosis factor and cycloheximide.  It can be
CC   used to produce Yama in a host cell for use in drug screening.
CC   Yama and anti-Yama protein nucleic acids, or antisense RNA, can be
CC   used to modulate cellular functions such as apoptosis and immune
CC   disorders mediated by the Fas pathway.
XX
SQ   Sequence 1158 BP; 328 A; 220 C; 275 G; 335 T; 0 other;


  Query Match             97.8%;  Score 1133.4;  DB 17;  Length 1158;
  Best Local Similarity   99.7%;  Pred. No. 4.1e-311;
  Matches 1156;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy       1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 gcacgagcggatgggtgctattgtgaggcggttgtagaagagtttcgtgagtgctcgcag 61

Qy      61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 ctcatacctgtggctgtgtatccgtggccacagctggttggcgtcgccttgaaatcccag 121

Qy     121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 gccgtgaggagttagcgagccctgctcacactcggcgctctggttttcggtgggtgtgcc 181

Qy     181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 ctgcacctgcctcttcccgcattctcattaataaaggtatccatggagaacactgaaaac 241

Qy     241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 tcagtggattcaaaatccattaaaaatttggaaccaaagatcatacatggaagcgaatca 301

Qy     301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 atggactctggaatatccctggacaacagttataaaatggattatcctgagatgggttta 361

Qy     361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 tgtataataattaataataagaattttcataaaagcactggaatgacatctcggtctggt 421

Qy     421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     422 acagatgtcgatgcagcaaacctcagggaaacattcagaaacttgaaatatgaagtcagg 481

Qy     481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     482 aataaaaatgatcttacacgtgaagaaattgtggaattgatgcgtgatgtttctaaagaa 541

Qy     541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     542 gatcacagcaaaaggagcagttttgtttgtgtgcttctgagccatggtgaagaaggaata 601

Qy     601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     602 atttttggaacaaatggacctgttgacctgaaaaaaataacaaacttttcagaggggat 661

Qy     661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     662 cgttgtagaagtctaactggaaaacccaaacttttcattattcaggcctgccgtggtaca 721

Qy     721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     722 gaactggactgtggcattgagacagacagtggtgttgatgatgacatggcgtgtcataaa 781

Qy     781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     782 ataccagtggaggccgacttcttgtatgcatactccacagcacctggttattattcttgg 841

Qy     841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     842 cgaaattcaaaggatggctcctggttcatccagtcgctttgtgccatgctgaaacagtat 901

Qy     901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     902 gccgacaagcttgaatttatgcacattcttacccgggttaaccgaaaggtggcaacagaa 961

Qy     961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     962 tttgagtcctttttcctttgacgctacttttcatgcaaagaaacagattccatgtattgtt 1021

Qy    1021 TCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1022 tccatgctcacaaaagaactctatttttatcactaaagaaatggttggttggtggttttt 1081

Qy    1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTG 1140
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    1082 tttagtttgtatgccaagtgagaagatggtatattt-ggtactgtatttccctctcattt 1140

Qy    1141 GGGACCTACTCTCATGCTG 1159
           ||| ||||||||||||||||
Db    1141 ggg-cctactctcatgctg 1158


RESULT   3
AAV05471
ID   AAV05471 standard; DNA; 1158 BP.
XX
AC   AAV05471;
XX
DT   02-JUL-1998  (first entry)
XX
DE   Nucleic acid encoding a protein designated YAMA.
XX
KW   CrmA; tumour necrosis factor receptor; TNF-R; U1-70; inhibition;
KW   apoptosis; activation; U1-70 path; apoptosis path; identification;
KW   modulation; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             224..1057
FT                   /*tag=  a
XX
PN   JP09299077-A.
XX
PD   25-NOV-1997.
XX
PF   27-JAN-1997;   97JP-0012932.
XX
PR   09-FEB-1996;   96US-0591605.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
DR   WPI; 1998-056551/06.
DR   P-PSDB; AAW41688.
XX
PT   Controlling apoptosis by inhibiting activation of the U1-70 pathway
PT   - by introduction of nucleic acid encoding protein with CrmA
PT   biological activity into the cell
XX
PS   Disclosure; Fig 1; 48pp; Japanese.
XX
CC   The present sequence encodes a non-naturally occurring protein designated
CC   YAMA. The specification describes a novel method of controlling cell
CC   functions controlled by the tumour necrosis factor receptor (TNF-R)
CC   pathway containing U1-70. The method comprises introducing a nucleic
CC   acid molecule encoding CrmA into a cell, and culturing the cell under
CC   suitable conditions for transcription and translation of the CrmA nucleic
```

```
CC   acid. Method for preventing and inhibiting apoptosis by inhibiting
CC   activation of the U1-70 path, identifying the chemical related to the
CC   apoptosis path, and screening a candidate chemical having biological
CC   function in the apoptosis path are also described. The methods can be
CC   used to identify substances which modulate apoptosis in cells controlled
CC   by the TNF-R pathway.
XX
SQ   Sequence 1158 BP; 328 A; 220 C; 275 G; 335 T; 0 other;


  Query Match             97.8%;  Score 1133.4;  DB 19;  Length 1158;
  Best Local Similarity   99.7%;  Pred. No. 4.1e-311;
  Matches 1156;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy       1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 gcacgagcggatgggtgctattgtgaggcggttgtagaagagtttcgtgagtgctcgcag 61

Qy      61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 ctcatacctgtggctgtgtatccgtggccacagctggttggcgtcgccttgaaatcccag 121

Qy     121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 gccgtgaggagttagcgagccctgctcacactcggcgctctggttttcggtgggtgtgcc 181

Qy     181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 ctgcacctgcctcttcccgcattctcattaataaaggtatccatggagaacactgaaaac 241

Qy     241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 tcagtggattcaaaatccattaaaaatttggaaccaaagatcatacatggaagcgaatca 301

Qy     301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 atggactctggaatatccctggacaacagttataaaatggattatcctgagatgggttta 361

Qy     361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 tgtataataattaataataagaattttcataaaagcactggaatgacatctcggtctggt 421

Qy     421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     422 acagatgtcgatgcagcaaacctcagggaaacattcagaaacttgaaatatgaagtcagg 481

Qy     481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     482 aataaaaatgatcttacacgtgaagaaattgtggaattgatgcgtgatgtttctaaagaa 541

Qy     541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     542 gatcacagcaaaaggagcagttttgtttgtgtgcttctgagccatggtgaagaaggaata 601

Qy     601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     602 atttttggaacaaatggacctgttgacctgaaaaaaataacaaacttttcagaggggat 661

Qy     661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     662 cgttgtagaagtctaactggaaaacccaaacttttcattattcaggcctgccgtggtaca 721

Qy     721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     722 gaactggactgtggcattgagacagacagtggtgttgatgatgacatggcgtgtcataaa 781

Qy     781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     782 ataccagtggaggccgacttcttgtatgcatactccacagcacctggttattattcttgg 841

Qy     841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     842 cgaaattcaaaggatggctcctggttcatccagtcgctttgtgccatgctgaaacagtat 901

Qy     901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     902 gccgacaagcttgaatttatgcacattcttacccgggttaaccgaaaggtggcaacagaa 961

Qy     961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     962 tttgagtccttttccttgacgctacttttcatgcaaagaaacagattccatgtattgtt 1021

Qy    1021 TCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1022 tccatgctcacaaaagaactctatttttatcactaaagaaatggttggttggtggttttt 1081

Qy    1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTG 1140
           |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    1082 tttagtttgtatgccaagtgagaagatggtatattt-ggtactgtatttccctctcattt 1140

Qy    1141 GGGACCTACTCTCATGCTG 1159
           ||| ||||||||||||||||
Db    1141 ggg-cctactctcatgctg 1158


RESULT   4
AAS10503
ID   AAS10503 standard; DNA; 2635 BP.
XX
AC   AAS10503;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Human caspase 3 DNA sequence.
XX
KW   Human; caspase 3; apoptosis; hyperproliferative disorder; hepatitis;
KW   viral infection; haematopoietic disorder; autoimmune disorder;
KW   atherosclerosis; neurological disorder; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             225..1058
FT                   /*tag= a
FT                   /product= "Caspase 3"
XX
PN   WO200153310-A1.
XX
PD   26-JUL-2001.
XX
PF   11-JAN-2001; 2001WO-US00888.
XX
PR   18-JAN-2000; 2000US-0484617.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Zhang H,  Cowsert LM;
XX
DR   WPI; 2001-442252/47.
DR   P-PSDB; AAU05394.
XX
PT   New antisense compound to inhibit caspase 3 is useful for treating
PT   hepatitis and atherosclerosis -
XX
PS   Example 13; Page 92-94; 127pp; English.
XX
CC   The present sequence encoding for human caspase 3 is described
CC   in an invention relating to novel antisense oligonucleotides
CC   (AAS10517-AAS10676) and methods of using these compounds for the
CC   modulation of caspase 3 expression. The caspase 3 antisense
CC   oligonucleotides specifically hybridise with and inhibit the
CC   expression of caspase 3. Antisense compounds targeted to caspase 3
CC   are useful to inhibit caspase 3 expression in cells or tissues and
CC   to modulate apoptosis. The caspase 3 antisense oligonucleotides are
CC   useful for treating disorders associated with expression of caspase 3.
CC   Such disorders include hyperproliferative disorders (e.g. cancer),
```

```
CC   viral infections (e.g. hepatitis), haematopoietic disorders, autoimmune
CC   disorders, atherosclerosis and neurological disorders (e.g. Alzheimer's
CC   disease).
XX
SQ   Sequence 2635 BP; 815 A; 465 C; 569 G; 786 T; 0 other;


  Query Match             96.5%;  Score 1118;  DB 22;  Length 2635;
  Best Local Similarity   99.4%;  Pred. No. 1.4e-306;
  Matches 1143;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    2;

Qy      10 GATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCT 69
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12 gaggggtgctattgtgaggcggttgtagaagagtttcgtgagtgctcgcagctcatacct 71

Qy      70 GTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72 gtggctgtgtatccgtggccacagctggttggcgtcgccttgaaatcccaggccgtgagg 131

Qy     130 AGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     132 agttagcgagccctgctcacactcggcgctctggttttcggtgggtgtgccctgcacctg 191

Qy     190 CCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGAT 249
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     192 cctcttcccccattctcattaataaaggtatccatggagaacactgaaaactcagtggat 251

Qy     250 TCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCT 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     252 tcaaaatccattaaaaatttggaaccaaagatcatacatggaagcgaatcaatggactct 311

Qy     310 GGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     312 ggaatatccctggacaacagttataaaatggattatcctgagatgggtttatgtataata 371

Qy     370 ATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTC 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     372 attaataataagaattttcataaaagcactggaatgacatctcggtctggtacagatgtc 431

Qy     430 GATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAAT 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     432 gatgcagcaaacctcagggaaacattcagaaacttgaaatatgaagtcaggaataaaaat 491

Qy     490 GATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     492 gatcttacacgtgaagaaattgtggaattgatgcgtgatgtttctaaagaagatcacagc 551

Qy     550 AAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGA 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     552 aaaaggagcagttttgtttgtgtgcttctgagccatggtgaagaaggaataatttttgga 611

Qy     610 ACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGA 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     612 acaaatggacctgttgacctgaaaaaaataacaaacttttcagaggggatcgttgtaga 671

Qy     670 AGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGAC 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     672 agtctaactggaaaacccaaacttttcattattcaggcctgccgtggtacagaactggac 731

Qy     730 TGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTG 789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     732 tgtggcattgagacagacagtggtgttgatgatgacatggcgtgtcataaaataccagtg 791

Qy     790 GAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCA 849
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     792 gatgccgacttcttgtatgcatactccacagcacctggttattattcttggcgaaattca 851

Qy     850 AAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAG 909
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     852 aaggatggctcctggttcatccagtcgctttgtgccatgctgaaacagtatgccgacaag 911

Qy     910 CTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCC 969
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db     912 cttgaatttatgcacattcttacccgggttaaccgaaaggtggcaacagaatttgagtcc 971

Qy     970 TTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTC 1029
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     972 ttttcctttgacgctacttttcatgcaaagaaacagattccatgtattgtttccatgctc 1031

Qy    1030 ACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTG 1089
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1032 acaaaagaactctatttttatcactaaagaaatggttggttggtggttttttttagtttg 1091

Qy    1090 TATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTAC 1149
           |||||||||||||||||||||||||| |||||||||||||||||||||||   |||||||
Db    1092 tatgccaagtgagaagatggtatattt-ggtactgtatttccctctcatt-ttgacctac 1149

Qy    1150 TCTCATGCTG 1159
           ||||||||||
Db    1150 tctcatgctg 1159


RESULT   5
AAD03910
ID   AAD03910 standard; cDNA; 2635 BP.
XX
AC   AAD03910;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Human caspase-3 cDNA.
XX
KW   Human; caspase-3; CPP32; yama; apopain;
KW   cysteine protease; apoptosis; caspase expression cassette; metastasis;
KW   tumour; cathepsin B; urokinase; proliferation; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             225..1058
FT                   /*tag= a
FT                   /product= "Human caspase-3"
XX
PN   WO200129232-A2.
XX
PD   26-APR-2001.
XX
PF   19-OCT-2000; 2000WO-US28941.
XX
PR   20-OCT-1999;   99US-0160559.
PR   14-AUG-2000; 2000US-0225564.
XX
PA   (SCIO-) SCIOS INC.
XX
PI   Cordell B,  Li Y;
XX
DR   WPI; 2001-290920/30.
DR   P-PSDB; AAE00600.
XX
PT   Novel fusion polypeptide comprising first and second caspase subunit
PT   separated by cleavage site not associated in nature with caspase
PT   subunit, useful for cloning gene encoding enzymes involved in
PT   proteolytic cleavage  -
XX
PS   Example 1; Fig 5; 116pp; English.
XX
CC   The present sequence is a cDNA encoding human Caspase-3 also known as
CC   CPP32, yama and apopain. Caspases are a family of cysteine proteases,
CC   that participate in the initiation and execution of apoptosis. Caspases
CC   exist as pro-enzymes, activated by cleavage into a large and small
CC   subunit, occurring after specific aspartic acid residues within the
CC   pro-enzyme sequence. The present invention relates to a method for
CC   functional cloning of genes encoding proteins or enzymes involved in
CC   proteolytic cleavage. The invention is based on the use of caspase
CC   expression cassettes comprising the coding sequence of a proteolytic
CC   cleavage site flanked by sequences encoding two caspase subunits. A
```

```
CC   fusion polypeptide comprising a first and a second caspase subunit,
CC   separated by a cleavage site not associated in nature, is useful for
CC   cloning gene encoding enzymes involved in proteolytic cleavage. An
CC   expression cassette containing fusion polypeptide is used to identify a
CC   mutant cell line deficient in an enzyme of interest and is also useful
CC   for diagnosis and suppression of proliferation or metastases of a tumour
CC   cell characterised by overexpression of a polypeptide (e.g. Cathepsin B
CC   or urokinase, selectively expressed in the tumour cells). DNA encoding
CC   fusion polypeptide is used in gene therapy.
XX
SQ   Sequence 2635 BP; 815 A; 465 C; 569 G; 786 T; 0 other;


  Query Match             96.5%;  Score 1118;  DB 22;  Length 2635;
  Best Local Similarity   99.4%;  Pred. No. 1.4e-306;
  Matches 1143;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    2;

Qy      10 GATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCT 69
           || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12 gaggggtgctattgtgaggcggttgtagaagagtttcgtgagtgctcgcagctcatacct 71

Qy      70 GTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72 gtggctgtgtatccgtggccacagctggttggcgtcgccttgaaatcccaggccgtgagg 131

Qy     130 AGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     132 agttagcgagccctgctcacactcggcgctctggttttcggtgggtgtgccctgcacctg 191

Qy     190 CCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGAT 249
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     192 cctcttccccattctcattaataaaggtatccatggagaacactgaaaactcagtggat 251

Qy     250 TCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCT 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     252 tcaaaatccattaaaaatttggaaccaaagatcatacatggaagcgaatcaatggactct 311

Qy     310 GGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     312 ggaatatccctggacaacagttataaaatggattatcctgagatgggtttatgtataata 371

Qy     370 ATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTC 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     372 attaataataagaattttcataaaagcactggaatgacatctcggtctggtacagatgtc 431

Qy     430 GATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAAT 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     432 gatgcagcaaacctcagggaaacattcagaaacttgaaatatgaagtcaggaataaaaat 491

Qy     490 GATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     492 gatcttacacgtgaagaaattgtggaattgatgcgtgatgtttctaaagaagatcacagc 551

Qy     550 AAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGA 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     552 aaaaggagcagttttgtttgtgtgcttctgagccatggtgaagaaggaataatttttgga 611

Qy     610 ACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGA 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     612 acaaatggacctgttgacctgaaaaaaataacaaacttttcagaggggatcgttgtaga 671

Qy     670 AGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGAC 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     672 agtctaactggaaaacccaaacttttcattattcaggcctgccgtggtacagaactggac 731

Qy     730 TGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTG 789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     732 tgtggcattgagacagacagtggtgttgatgatgacatggcgtgtcataaaatacCagtg 791

Qy     790 GAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCA 849
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     792 gatgccgacttcttgtatgcatactccacagcacctggttattattcttggcgaaattca 851

Qy     850 AAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAG 909
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     852 aaggatggctcctggttcatccagtcgctttgtgccatgctgaaacagtatgccgacaag 911

Qy     910 CTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCC 969
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     912 cttgaatttatgcacattcttacccgggttaaccgaaaggtggcaacagaatttgagtcc 971

Qy     970 TTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTC 1029
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     972 ttttcctttgacgctacttttcatgcaaagaaacagattccatgtattgtttccatgctc 1031

Qy    1030 ACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTG 1089
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1032 acaaaagaactctatttttatcactaaagaaatggttggttggtggttttttttagtttg 1091

Qy    1090 TATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTAC 1149
           |||||||||||||||||||||||||| |||||||||||||||||||||   |||||||
Db    1092 tatgccaagtgagaagatggtatattt-ggtactgtatttccctctcatt-ttgacctac 1149

Qy    1150 TCTCATGCTG 1159
           ||||||||||
Db    1150 tctcatgctg 1159


RESULT   6
AAV32608
ID   AAV32608 standard; cDNA; 1001 BP.
XX
AC   AAV32608;
XX
DT   26-OCT-1998  (first entry)
XX
DE   Mutant human apopain (caspase-3) pro-enzyme cDNA.
XX
KW   Mutant human apopain pro-enzyme; caspase-3; protease; SPA;
KW   scintillation proximity binding assay; diabetes; cancer;
KW   4-phosphono(difluoromethyl) phenylalanine group; osteoporosis;
KW   tyrosine phosphatase; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             64..897
FT                   /*tag= a
FT                   /product= "Mutant apopain C163S pro-enzyme"
FT   misc_feature    148..588
FT                   /*tag= b
FT                   /note= "Encodes the P17 subunit"
FT   mutation        550..552
FT                   /*tag= c
FT                   /note= "Changed from TGC in wild-type to TCC in
FT                        mutant"
FT   misc_feature    589..894
FT                   /*tag= d
FT                   /note= "Encodes the P12 subunit"
XX
PN   WO9820024-A1.
XX
PD   14-MAY-1998.
XX
PF   03-NOV-1997;   97WO-CA00824.
XX
PR   04-NOV-1996;   96US-0030411.
XX
PA   (MERI ) MERCK FROSST CANADA INC.
XX
PI   Desmarais S,  Friesen R,  Zamboni R;
XX
DR   WPI; 1998-348101/30.
DR   P-PSDB; AAW48937.
XX
```

```
PT   Peptide(s) useful in binding assays for tyrosine phosphatases or
PT   cysteine proteases - contain two or more 4-phosphono(difluoromethyl)
PT   phenylalanine groups to improve binding affinity
XX
PS   Disclosure; Pages 38-39; 59pp; English.
XX
CC   The present sequence represents a mutated human apopain (caspase-3)
CC   pro-enzyme cDNA.  The cDNA encodes a mutant human apopain C163S
CC   pro-enzyme.  The invention provides a method for use in a
CC   scintillation proximity binding assay (SPA) for proteases and
CC   phosphatases.  The method involves using mutated proteases and
CC   phosphatases whereby the catalytic cysteine residue of the enzymes
CC   are replaced with a serine or alanine residue to correct the problem of
CC   interference in SPA from extraneous oxidising and alkylating agents.
CC   The mutation affects the catalytic properties of the enzyme but does
CC   not affect their binding properties.  The invention claims for new
CC   ligands for use in SPA which have increased binding affinity for a
CC   tyrosine phosphatase or cysteine protease.  The ligands contain at least
CC   two 4-phosphono(difluoromethyl) phenylalanine groups which increase
CC   binding affinity of the ligand to its respective enzyme.  The assay can
CC   be used to determine the ability of new ligands and compound mixtures to
CC   competitively bind with an enzyme.  The method is claimed to allow
CC   a better usage of SPA in the discovery of compounds for the treatment
CC   and study of diseases, e.g. diabetes, cancer and osteoporosis.
XX
SQ   Sequence 1001 BP; 310 A; 172 C; 226 G; 293 T; 0 other;

  Query Match           79.5%;  Score 921.6;  DB 19;  Length 1001;
  Best Local Similarity  99.4%;  Pred. No. 4.5e-251;
  Matches  946;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 ttaataaaggtatccatggagaacactgaaaactcagtggattcaaaatccattaaaaat 108

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 ttggaaccaaagatcatacatggaagcgaatcaatggactctggaatatccctggacaac 168

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 agttataaaatggattatcctgagatgggtttatgtataataattaataataagaatttt 228

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 cataagagcactggaatgacatctcggtctggtacagatgtcgatgcagcaaacctcagg 288

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 gaaacattcagaaacttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaa 348

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 attgtggaattgatgcgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtt 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 tgtgtgcttctgagccatggtgaagaaggaataatttttggaacaaatggacctgttgac 468

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 ctgaaaaaaataacaaacttttcagaggggatcgttgtagaagtctaactggaaaaccc 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    529 aaacttttcattattcaggcctcccgtggtacagaactggactgtggcattgagacagac 588

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 agtggtgttgatgatgacatggcgtgtcataaaatataccagtggaggccgacttcttgtat 648

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 gcatactccacagcacctggttattattcttggcgaaattcaaaggatggctcctggttc 708

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    709 atccagtcgctttgtgccatgctgaaacagtatgccgacaagcttgaatttatgcacatt 768

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 cttacccgggttaaccgaaaggtggcaacagaatttgagtccttttccttttgacgctact 828

Qy    988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 tttcatgcaaagaaacagattccatgtattgtttccatgctcacaaaagaactctatttt 888

Qy   1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 tatcactaaagaaatggttggttggtggttttttttagtttgtatgccaagtgagaagat 948

Qy   1108 GGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTACTCTCATGCTG 1159
          ||||||||| |||||||||||||||||||||   ||||||||||||||||
Db    949 ggtatattt-ggtactgtatttccctctcatt-ttgacctactctcatgctg 998


RESULT   7
AAV32615
ID   AAV32615 standard; cDNA; 1001 BP.
XX
AC   AAV32615;
XX
DT   26-OCT-1998  (first entry)
XX
DE   Mutant human apopain (caspase-3) pro-enzyme cDNA.
XX
KW   Mutant human apopain pro-enzyme; caspase-3; protease; SPA;
KW   scintillation proximity binding assay; diabetes; cancer;
KW   4-phosphono(difluoromethyl) phenylalanine group; osteoporosis;
KW   tyrosine phosphatase; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             64..897
FT                   /*tag= a
FT                   /product= "Mutant apopain C163S pro-enzyme"
FT   misc_feature    148..588
FT                   /*tag= b
FT                   /note= "Encodes the P17 subunit"
FT   mutation        550..552
FT                   /*tag= c
FT                   /note= "Changed from TGC in wild-type to TCC in
FT                   mutant"
FT   misc_feature    589..894
FT                   /*tag= d
FT                   /note= "Encodes the P12 subunit"
XX
PN   WO9820156-A1.
XX
PD   14-MAY-1998.
XX
PF   03-NOV-1997;   97WO-CA00825.
XX
PR   04-NOV-1996;   96US-0030408.
XX
PA   (MERI ) MERCK FROSST CANADA INC.
XX
PI   Desmarais S,  Ford-Hutchinson A,  Friesen R,  Gresser M;
PI   Kennedy B,  Nicholson D,  Ramachandran C,  Skorey K;
XX
DR   WPI; 1998-348103/30.
XX
PT   Enzyme binding assay for detection of useful compounds - uses mutant
```

PT form of wild-type enzyme, in which serine replaces cysteine at
PT active site, to reduce interference from oxidising/alkylating agents
XX
PS Disclosure; Pages 38-39; 63pp; English.
XX
CC The present sequence represents a mutated human apopain (caspase-3)
CC pro-enzyme cDNA. The cDNA encodes a mutant human apopain C163S
CC pro-enzyme. The invention claims for a method for use in a
CC scintillation proximity binding assay (SPA) for proteases and
CC phosphatases. The method involves using mutated proteases and
CC phosphatases whereby the catalytic cysteine residue of the enzymes
CC are replaced with a serine or alanine residue to correct the problem of
CC interference in SPA from extraneous oxidising and alkylating agents.
CC The mutation affects the catalytic properties of the enzyme but does
CC not affect their binding properties. The invention provides new
CC ligands for use in SPA which have increased binding affinity for a
CC tyrosine phosphatase or cysteine protease. The ligands contain at least
CC two 4-phosphono(difluoromethyl) phenylalanine groups which increase
CC binding affinity of the ligand to its respective enzyme. The assay can
CC be used to determine the ability of new ligands and compound mixtures to
CC competitively bind with an enzyme. The method is claimed to allow
CC a better usage of SPA in the discovery of compounds for the treatment
CC and study of diseases, e.g. diabetes, cancer and osteoporosis.
XX
SQ Sequence 1001 BP; 310 A; 172 C; 226 G; 293 T; 0 other;

Query Match 79.5%; Score 921.6; DB 19; Length 1001;
Best Local Similarity 99.4%; Pred. No. 4.5e-251;
Matches 946; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```
Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 ttaataaaggtatccatggagaacactgaaaactcagtggattcaaaatccattaaaaat 108

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 ttggaaccaaagatcatacatggaagcgaatcaatggactctggaatatccctggacaac 168

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 agttataaaatggattatcctgagatgggtttatgtataataattaataataagaatttt 228

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 cataagagcactggaatgacatctcggtctggtacagatgtcgatgcagcaaacctcagg 288

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 gaaacattcagaaacttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaa 348

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 attgtggaattgatgcgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtt 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 tgtgtgcttctgagccatggtgaagaaggaataattttttggaacaaatggacctgttgac 468

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 ctgaaaaaaataacaaacttttttcagaggggatcgttgtagaagtctaactggaaaaccc 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    529 aaacttttcattattcaggcctcccgtggtacagaactggactgtggcattgagacagac 588

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 agtggtgttgatgatgacatggcgtgtcataaaatacccagtggaggccgacttcttgtat 648

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 gcatactccacagcacctggttattattcttggcgaaattcaaaggatggctcctggttc 708

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 atccagtcgctttgtgccatgctgaaacagtatgccgacaagcttgaatttatgcacatt 768

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 cttacccgggttaaccgaaaggtggcaacagaatttgagtccttttccttgacgctact 828

Qy    988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 tttcatgcaaagaaacagattccatgtattgtttccatgctcacaaaagaactctatttt 888

Qy   1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 tatcactaaagaaatggttggttggtggttttttttagtttgtatgccaagtgagaagat 948

Qy   1108 GGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTACTCTCATGCTG 1159
          ||||||||| |||||||||||||||||||||   |||||||||||||||||
Db    949 ggtatattt-ggtactgtatttccctctcatt-ttgacctactctcatgctg 998
```

RESULT 8
AAC64043
ID AAC64043 standard; DNA; 2538 BP.
XX
AC AAC64043;
XX
DT 14-FEB-2001 (first entry)
XX
DE pYNGCas3p10p35 insert encoding human caspase-3 and AcNPV p35 protein.
XX
KW Human caspase-3; AcNPV p35 protein; BmNPV p10 promoter;
KW polyhedron promoter; baculovirus vector; apoptosis inhibitor protein;
KW recominant caspase production; insect cell; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Autographa californica nuclear polyhedrosis virus.
OS Chimeric - Bombyx mori nuclear polyhedrosis virus.
XX
PN JP2000253882-A.
XX
PD 19-SEP-2000.
XX
PF 10-MAR-1999; 99JP-0062732.
XX
PR 10-MAR-1999; 99JP-0062732.
XX
PA (KATA) KATAKURA IND CO LTD.
XX
DR WPI; 2000-658811/64.
XX
PT Baculovirus vector used for the preparation of caspase protein,
PT comprises a transfer vector plasmid containing a caspase
PT protein-producing gene -
XX
PS Example 1; Page 7-8; 10pp; Japanese.
XX
CC The invention relates to a novel baculovirus vector used for the
CC preparation of a caspase protein, and containing a recombinant transfer
CC vector plasmid comprising a gene encoding a caspase and a gene encoding
CC an apoptosis-inhibiting protein. The invention also encompasses the
CC preparation of caspase protein, comprising transfecting the novel
CC baculovirus vector into an insect or an insect cell; breeding the
CC insect or culturing the cell; and collecting the caspase protein from
CC the insect or insect cell. The present sequence represents the expression
CC cassette insert from a transfer vector (pYNGCas3p10p35) produced in an
CC exemplification of the invention. This insert comprises a mature human
CC caspase-3 gene under the control of a polyhedron promoter, and an
CC Autographa californica nuclear polyhedrosis virus (AcNPV) p35 gene under
CC the control of a Bombyx mori nuclear polyhedrosis virus (BmNPV) p10
CC promoter.

```
XX
SQ   Sequence 2538 BP; 865 A; 446 C; 506 G; 721 T; 0 other;

  Query Match             72.4%;  Score 839.6;  DB 21;  Length 2538;
  Best Local Similarity   99.5%;  Pred. No. 1.3e-227;
  Matches  842;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy    211 ATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTG 270
          |||||   ||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 ataaatagatctatggagaacactgaaaactcagtggattcaaaatccattaaaaatttg 183

Qy    271 GAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGT 330
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 gaaccaaagatcatacatggaagcgaatcaatggactctggaatatccctggacaacagt 243

Qy    331 TATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCAT 390
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 tataaaatggattatcctgagatgggtttatgtataataattaataataagaattttcat 303

Qy    391 AAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAA 450
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 aaaagcactggaatgacatctcggtctggtacagatgtcgatgcagcaaacctcagggaa 363

Qy    451 ACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATT 510
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 acattcagaaacttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaaatt 423

Qy    511 GTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGT 570
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    424 gtggaattgatgcgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtttgt 483

Qy    571 GTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTG 630
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 gtgcttctgagccatggtgaagaaggaataatttttggaacaaatggacctgttgacctg 543

Qy    631 AAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAA 690
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544 aaaaaaataacaaacttttcagaggggatcgttgtagaagtctaactggaaaacccaaa 603

Qy    691 CTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGT 750
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604 cttttcattattcaggcctgccgtggtacagaactggactgtggcattgagacagacagt 663

Qy    751 GGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCA 810
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664 ggtgttgatgatgacatggcgtgtcataaaatacctgtggaggccgacttcttgtatgca 723

Qy    811 TACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATC 870
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    724 tactccacagcacctggttattattcttggcgaaattcaaaggatggctcctggttcatc 783

Qy    871 CAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTT 930
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    784 cagtcgctttgtgccatgctgaaacagtatgccgacaagcttgaatttatgcacattctt 843

Qy    931 ACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTT 990
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    844 acccgggttaaccgaaaggtggcaacagaatttgagtcctttttcctttgacgctacttttt 903

Qy    991 CATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTAT 1050
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    904 catgcaaagaaacagattccatgtattgtttccatgctcacaaaagaactctatttttat 963

Qy   1051 CACTAA 1056
          ||||||
Db    964 cactaa 969


RESULT   9
AAD13075
ID   AAD13075 standard; cDNA; 834 BP.
XX
AC   AAD13075;
XX
DT   16-OCT-2001  (first entry)
XX
DE   Mammalia caspase-3 reporter moiety cDNA.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance;
KW   caspase-3; ss.
XX
OS   Mammalia.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain -
XX
PS   Example 4; Page 131; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is mammalia caspase-3 cDNA which is used as a reporter moiety in
CC   the exemplification of the invention.
XX
SQ   Sequence 834 BP; 274 A; 147 C; 179 G; 234 T; 0 other;

  Query Match             72.0%;  Score 834;  DB 22;  Length 834;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-226;
  Matches  834;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    223 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATC 282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atggagaacactgaaaactcagtggattcaaaatccattaaaaatttggaaccaaagatc 60

Qy    283 ATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGAT 342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 atacatggaagcgaatcaatggactctggaatatccctggacaacagttataaaatggat 120

Qy    343 TATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGA 402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 tatcctgagatgggtttatgtataataattaataataagaattttcataaaagcactgga 180

Qy    403 ATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAAC 462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 atgacatctcggtctggtacagatgtcgatgcagcaaacctcagggaaacattcagaaac 240
```

```
Qy    463 TTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaaattgtggaattgatg 300

Qy    523 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGC 582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 cgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtttgtgtgcttctgagc 360

Qy    583 CATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACA 642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 catggtgaagaaggaataatttttggaacaaatggacctgttgacctgaaaaaaataaca 420

Qy    643 AACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATT 702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 aacttttcagaggggatcgttgtagaagtctaactggaaaacccaaactttcattatt 480

Qy    703 CAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGAT 762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 caggcctgccgtggtacagaactggactgtggcattgagacagacagtggtgttgatgat 540

Qy    763 GACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 gacatggcgtgtcataaaatacctgtggaggccgacttcttgtatgcatactccacagca 600

Qy    823 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGT 882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 cctggttattattcttggcgaaattcaaaggatggctcctggttcatccagtcgctttgt 660

Qy    883 GCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAAC 942
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 gccatgctgaaacagtatgccgacaagcttgaatttatgcacattcttacccgggttaac 720

Qy    943 CGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAA 1002
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 cgaaaggtggcaacagaatttgagtcctttttcctttgacgctacttttcatgcaaagaaa 780

Qy   1003 CAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAA 1056
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 cagattccatgtattgtttccatgctcacaaaagaactctatttttatcactaa 834
```

RESULT 10
AAX81219
ID AAX81219 standard; DNA; 835 BP.
XX
AC AAX81219;
XX
DT 10-SEP-1999 (first entry)
XX
DE Nucleotide sequence of caspase-3 (CPP32).
XX
KW Rev-caspase; cysteine protease; zymogen; caspase; cancer; gene therapy;
KW autoimmune disease; caspase-mediated apoptosis; neurodegenerative;
KW tumour cell; myocardial infarction; human; ds.
XX
OS Homo sapiens.
XX
PN WO9935277-A2.
XX
PD 15-JUL-1999.
XX
PF 11-JAN-1999; 99WO-US00632.
XX
PR 09-JAN-1998; 98US-0070987.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Alnemri ES;
XX
DR WPI; 1999-419353/35.
DR P-PSDB; AAY21717.
XX
PT New isolated nucleic acid molecule encoding a rev-caspase - used
PT for screening and identifying inhibitors or enhancers for treating
PT cancer or autoimmune disease
XX
PS Disclosure; Fig 13A-B; 74pp; English.
XX
CC The invention relates to nucleic acid molecules encoding rev-caspases.
CC Rev-caspases are cysteine proteases that specifically cleave proteins
CC after Asp residues and is expressed as a zymogen, in which the small
CC subunit is N-terminal to a large subunit. A gene delivery vehicle
CC comprising a rev-caspase coding sequence is useful for the treatment of
CC cancer, where the gene delivery vehicle is internalised by tumour cells.
CC The gene delivery vehicle can also be used to treat autoimmune diseases.
CC Cells transfected with a rev-caspase expressing vector can be used in
CC identification of inhibitors or enhancers of caspase-mediated apoptosis.
CC In vitro translated rev-caspase can be used to identify an inhibitor or
CC enhancer of caspase processing activity. Caspase inhibitors are useful
CC for treating neurodegenerative diseases as well as for inhibiting
CC apoptosis in the heart following myocardial infarction. Sequences
CC AAX81217 -AAX81226 represent human caspase genes encoding caspase 1-10
CC gene products (AAY21715-Y21724).
XX
SQ Sequence 835 BP; 274 A; 147 C; 178 G; 235 T; 1 other;

Query Match 71.8%; Score 832.4; DB 20; Length 835;
Best Local Similarity 99.9%; Pred. No. 8.1e-226;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy    223 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATC 282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atggagaacactgaaaactcagtggattcaaaatccattaaaaatttggaaccaaagatc 60

Qy    283 ATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGAT 342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 atacatggaagcgaatcaatggactctggaatatccctggacaacagttataaaatggat 120

Qy    343 TATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGA 402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 tatcctgagatgggtttatgtataataattaataataagaattttcataaaagcactgga 180

Qy    403 ATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAAC 462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 atgacatctcggtctggtacagatgtcgatgcagcaaacctcagggaaacattcagaaac 240

Qy    463 TTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaaattgtggaattgatg 300

Qy    523 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGC 582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 cgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtttgtgtgcttctgagc 360

Qy    583 CATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACA 642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 catggtgaagaaggaataatttttggaacaaatggacctgttgacctgaaaaaaataaca 420

Qy    643 AACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATT 702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 aacttttcagaggggatcgttgtagaagtctaactggaaaacccaaactttcattatt 480

Qy    703 CAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGAT 762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 caggcctgccgtggtacagaactggactgtggcattgagacagacagtggtgttgatgat 540

Qy    763 GACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 gacatggcgtgtcataaaatacctgtggaggccgacttcttgtatgcatactccacagca 600

Qy    823 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGT 882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 cctggttattattcttggcgaaattcaaaggatggctcctggttcatccagtcgctttgt 660
```

```
Qy    883 GCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAAC 942
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 gccatgctgaaacagtatgccgacaagcttgaatttatgcacattcttacccgggttaac 720

Qy    943 CGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAA 1002
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 cgaaaggtggcaacagaatttgagtccttttccttgacgctacttttcatgcaaagaaa 780

Qy   1003 CAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAA 1056
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 cagattccatgtattgtttccatgctcacaaaagaactctatttttatcactaa 834
```

RESULT 11
AAD13084
ID AAD13084 standard; DNA; 7247 BP.
XX
AC AAD13084;
XX
DT 16-OCT-2001 (first entry)
XX
DE pcDNA3-Ub-Met-C3 construct.
XX
KW Protein destabilisation; linker moiety; reporter moiety; disease model;
KW linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW transgenic animal; transgenic plant; disease resistance; caspase-3;
KW human; ubiquitin; cyclic; circular; plasmid pcDNA3-Ub-Met-C3;
KW fusion protein; ds.
XX
OS Chimeric - Mammalia.
OS Chimeric - Homo sapiens.
XX
PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US03791.
XX
PR 04-FEB-2000; 2000US-0498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by
PT -NH-ubiquitin protein endoproteases, comprising two copies of the
PT domain -
XX
PS Example 4; Page 140-141; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method
CC is also useful for developing novel assays for a wide range of
CC post-translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a pcDNA3-Ub-Met-C3 construct comprising wildtype multimerised
CC human ubiquitin destabilisation domain-Mammalia caspase-3 fusion protein.
XX
SQ Sequence 7247 BP; 1794 A; 1796 C; 1826 G; 1831 T; 0 other;

Query Match 71.3%; Score 826.4; DB 22; Length 7247;
Best Local Similarity 94.9%; Pred. No. 1.2e-223;
Matches 854; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

```
Qy    172 GGGTGTGCCCTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAAC 231
          | ||   ||||||||||   || |             ||  | | ||||      |||
Db   1093 gagtccaccctgcacctggtactccgtctcagaggtgggatgcacggatcccacatcaac 1152

Qy    232 ACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGA 291
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   1153 actgaaaactcagtggcctcaaaatccattaaaaatttggaaccaaagatcatacatgga 1212

Qy    292 AGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAG 351
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1213 agcgaatcaatggactctggaatatccctggacaacagttataaaatggattatcctgag 1272

Qy    352 ATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCT 411
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1273 atgggtttatgtataataattaataataagaattttcataaaagcactggaatgacatct 1332

Qy    412 CGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATAT 471
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1333 cggtctggtacagatgtcgatgcagcaaacctcagggaaacattcagaaacttgaaatat 1392

Qy    472 GAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTT 531
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1393 gaagtcaggaataaaaatgatcttacacgtgaagaaattgtggaattgatgcgtgatgtt 1452

Qy    532 TCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAA 591
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1453 tctaaagaagatcacagcaaaaggagcagttttgtttgtgtgcttctgagccatggtgaa 1512

Qy    592 GAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTC 651
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1513 gaaggaataatttttggaacaaatggacctgttgacctgaaaaaaataacaaacttttttc 1572

Qy    652 AGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGC 711
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1573 agaggggatcgttgtagaagtctaactggaaaacccaaacttttcattattcaggcctgc 1632

Qy    712 CGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCG 771
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1633 cgtggtacagaactggactgtggcattgagacagacagtggtgttgatgatgacatggcg 1692

Qy    772 TGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTAT 831
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   1693 tgtcataaaatacagtggatgccgacttcttgtatgcatactccacagcacctggttat 1752

Qy    832 TATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTG 891
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1753 tattcttggcgaaattcaaaggatggctcctggttcatccagtcgctttgtgccatgctg 1812

Qy    892 AAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTG 951
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1813 aaacagtatgccgacaagcttgaatttatgcacattcttacccgggttaaccgaaaggtg 1872

Qy    952 GCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCA 1011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1873 gcaacagaatttgagtcctttttcctttgacgctacttttcatgcaaagaaacagattcca 1932

Qy   1012 TGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTG 1071
          |||||||||||||||||||||||||||||||||||||  |     |  ||| ||
Db   1933 tgtattgtttccatgctcacaaaagaactctatttttatcacggatccggggcgtggctg 1992
```

RESULT 12
AAD13078
ID AAD13078 standard; DNA; 6436 BP.
XX
AC AAD13078;
XX
DT 16-OCT-2001 (first entry)
XX

```
DE   pcDNA3-1XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-1XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain  -
XX
PS   Example 4; Page 133-134; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-1XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domain-Mammalia caspase-3 fusion protein.
XX
SQ   Sequence 6436 BP; 1577 A; 1606 C; 1607 G; 1646 T; 0 other;

  Query Match             71.2%;  Score 825;  DB 22;  Length 6436;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-223;
  Matches  825;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     229 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1144 aacactgaaaactcagtggattcaaaatccattaaaaatttggaaccaaagatcatacat 1203

Qy     289 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 348
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1204 ggaagcgaatcaatggactctggaatatccctggacaacagttataaaatggattatcct 1263

Qy     349 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 408
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1264 gagatgggtttatgtataataattaataataagaattttcataaaagcactggaatgaca 1323

Qy     409 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 468
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1324 tctcggtctggtacagatgtcgatgcagcaaacctcagggaaacattcagaaacttgaaa 1383

Qy     469 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 528
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1384 tatgaagtcaggaataaaaatgatcttacacgtgaagaaattgtggaattgatgcgtgat 1443

Qy     529 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 588
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1444 gtttctaaagaagatcacagcaaaaggagcagttttgtttgtgtgcttctgagccatggt 1503

Qy     589 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 648
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1504 gaagaaggaataatttttggaacaaatggacctgttgacctgaaaaaaataacaaacttt 1563

Qy     649 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 708
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1564 ttcagaggggatcgttgtagaagtctaactggaaaacccaaacttttcattattcaggcc 1623

Qy     709 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 768
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1624 tgccgtggtacagaactggactgtggcattgagacagacagtggtgttgatgatgacatg 1683

Qy     769 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 828
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1684 gcgtgtcataaaatacCagtggaggccgacttcttgtatgcatactccacagcacctggt 1743

Qy     829 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 888
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1744 tattattcttggcgaaattcaaaggatggctcctggttcatccagtcgctttgtgccatg 1803

Qy     889 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1804 ctgaaacagtatgccgacaagcttgaatttatgcacattcttacccgggttaaccgaaag 1863

Qy     949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1008
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1864 gtggcaacagaatttgagtccttttccttgacgctacttttcatgcaaagaaacagatt 1923

Qy    1009 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1053
           |||||||||||||||||||||||||||||||||||||||||||||
Db    1924 ccatgtattgtttccatgctcacaaaagaactctatttttatcac 1968


RESULT  13
AAD13079
ID   AAD13079 standard; DNA; 6703 BP.
XX
AC   AAD13079;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-2XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-2XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
```

PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by
PT -NH-ubiquitin protein endoproteases, comprising two copies of the
PT domain -
XX
PS Example 4; Page 134-135; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method
CC is also useful for developing novel assays for a wide range of
CC post-translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a pcDNA3-2XUb-C3 construct comprising mutant multimerised
CC human ubiquitin destabilisation domains-Mammalia caspase-3
CC fusion protein.
XX
SQ Sequence 6703 BP; 1653 A; 1677 C; 1674 G; 1699 T; 0 other;

Query Match 71.2%; Score 825; DB 22; Length 6703;
Best Local Similarity 100.0%; Pred. No. 2.8e-223;
Matches 825; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    229 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 288
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1411 aacactgaaaactcagtggattcaaaatccattaaaaatttggaaccaaagatcatacat 1470

Qy    289 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 348
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1471 ggaagcgaatcaatggactctggaatatccctggacaacagttataaaatggattatcct 1530

Qy    349 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 408
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1531 gagatgggtttatgtataataattaataataagaattttcataaaagcactggaatgaca 1590

Qy    409 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 468
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1591 tctcggtctggtacagatgtcgatgcagcaaacctcagggaaacattcagaaacttgaaa 1650

Qy    469 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 528
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1651 tatgaagtcaggaataaaaatgatcttacacgtgaagaaattgtggaattgatgcgtgat 1710

Qy    529 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 588
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1711 gtttctaaagaagatcacagcaaaaggagcagttttgtttgtgtgcttctgagccatggt 1770

Qy    589 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 648
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1771 gaagaaggaataatttttggaacaaatggacctgttgacctgaaaaaaataacaaacttt 1830

Qy    649 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 708
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1831 ttcagaggggatcgttgtagaagtctaactggaaaacccaaacttttcattattcaggcc 1890

Qy    709 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 768
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1891 tgccgtggtacagaactggactgtggcattgagacagacagtggtgttgatgatgacatg 1950

Qy    769 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 828
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1951 gcgtgtcataaaataccagtggaggccgacttcttgtatgcatactccacagcacctggt 2010

Qy    829 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 888
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2011 tattattcttggcgaaattcaaaggatggctcctggttcatccagtcgctttgtgccatg 2070

Qy    889 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2071 ctgaaacagtatgccgacaagcttgaatttatgcacattcttacccgggttaaccgaaag 2130

Qy    949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1008
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2131 gtggcaacagaatttgagtccttttccttgacgctacttttcatgcaaagaaacagatt 2190

Qy   1009 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1053
          |||||||||||||||||||||||||||||||||||||||||||||
Db   2191 ccatgtattgtttccatgctcacaaaagaactctatttttatcac 2235
```

RESULT 14
AAD13080
ID AAD13080 standard; DNA; 6946 BP.
XX
AC AAD13080;
XX
DT 16-OCT-2001 (first entry)
XX
DE pcDNA3-3XUb-C3 construct.
XX
KW Protein destabilisation; linker moiety; reporter moiety; disease model;
KW linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW transgenic animal; transgenic plant; disease resistance; caspase-3;
KW human; ubiquitin; cyclic; circular; plasmid pcDNA3-3XUb-C3;
KW fusion protein; ds.
XX
OS Chimeric - Mammalia.
OS Chimeric - Homo sapiens.
XX
PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US03791.
XX
PR 04-FEB-2000; 2000US-0498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by
PT -NH-ubiquitin protein endoproteases, comprising two copies of the
PT domain -
XX
PS Example 4; Page 136-137; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method
CC is also useful for developing novel assays for a wide range of
CC post-translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a pcDNA3-3XUb-C3 construct comprising mutant multimerised
CC human ubiquitin destabilisation domains-Mammalia caspase-3
CC fusion protein.

```
XX
SQ   Sequence 6946 BP; 1723 A; 1742 C; 1735 G; 1746 T; 0 other;

  Query Match           71.2%;  Score 825;  DB 22;  Length 6946;
  Best Local Similarity  100.0%;  Pred. No. 2.8e-223;
  Matches  825;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     229 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1654 aacactgaaaactcagtggattcaaaatccattaaaaatttggaaccaaagatcatacat 1713

Qy     289 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 348
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1714 ggaagcgaatcaatggactctggaatatccctggacaacagttataaaatggattatcct 1773

Qy     349 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 408
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1774 gagatgggtttatgtataataattaataataagaattttcataaaagcactggaatgaca 1833

Qy     409 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 468
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1834 tctcggtctggtacagatgtcgatgcagcaaacctcagggaaacattcagaaacttgaaa 1893

Qy     469 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 528
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1894 tatgaagtcaggaataaaaatgatcttacacgtgaagaaattgtggaattgatgcgtgat 1953

Qy     529 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 588
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1954 gtttctaaagaagatcacagcaaaaggagcagttttgtttgtgtgcttctgagccatggt 2013

Qy     589 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 648
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2014 gaagaaggaataatttttggaacaaatggacctgttgacctgaaaaaaataacaaacttt 2073

Qy     649 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 708
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2074 ttcagaggggatcgttgtagaagtctaactggaaaacccaaacttttcattattcaggcc 2133

Qy     709 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 768
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2134 tgccgtggtacagaactggactgtggcattgagacagacagtggtgttgatgatgacatg 2193

Qy     769 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 828
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2194 gcgtgtcataaaatacccagtggaggccgacttcttgtatgcatactccacagcacctggt 2253

Qy     829 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 888
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2254 tattattcttggcgaaattcaaaggatggctcctggttcatccagtcgctttgtgccatg 2313

Qy     889 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2314 ctgaaacagtatgccgacaagcttgaatttatgcacattcttacccgggttaaccgaaag 2373

Qy     949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1008
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2374 gtggcaacagaatttgagtccttttccttgacgctacttttcatgcaaagaaacagatt 2433

Qy    1009 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1053
           |||||||||||||||||||||||||||||||||||||||||||||
Db    2434 ccatgtattgtttccatgctcacaaaagaactctatttttatcac 2478


RESULT  15
AAD13081
ID   AAD13081 standard; DNA; 7189 BP.
XX
AC   AAD13081;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-4XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-4XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain -
XX
PS   Example 4; Page 137-138; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabillsing domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-4XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domains-Mammalia caspase-3
CC   fusion protein.
XX
SQ   Sequence 7189 BP; 1793 A; 1807 C; 1796 G; 1793 T; 0 other;

  Query Match           71.2%;  Score 825;  DB 22;  Length 7189;
  Best Local Similarity  100.0%;  Pred. No. 2.9e-223;
  Matches  825;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     229 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1897 aacactgaaaactcagtggattcaaaatccattaaaaatttggaaccaaagatcatacat 1956

Qy     289 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 348
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1957 ggaagcgaatcaatggactctggaatatccctggacaacagttataaaatggattatcct 2016

Qy     349 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 408
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2017 gagatgggtttatgtataataattaataataagaattttcataaaagcactggaatgaca 2076

Qy     409 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 468
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2077 tctcggtctggtacagatgtcgatgcagcaaacctcagggaaacattcagaaacttgaaa 2136
```

```
Qy     469 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 528
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2137 tatgaagtcaggaataaaaatgatcttacacgtgaagaaattgtggaattgatgcgtgat 2196

Qy     529 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 588
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2197 gtttctaaagaagatcacagcaaaaggagcagttttgtttgtgtgcttctgagccatggt 2256

Qy     589 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 648
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2257 gaagaaggaataatttttggaacaaatggacctgttgacctgaaaaaaataacaaacttt 2316

Qy     649 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 708
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2317 ttcagaggggatcgttgtagaagtctaactggaaaacccaaacttttcattattcaggcc 2376

Qy     709 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 768
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2377 tgccgtggtacagaactggactgtggcattgagacagacagtggtgttgatgatgacatg 2436

Qy     769 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 828
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2437 gcgtgtcataaaataccagtggaggccgacttcttgtatgcatactccacagcacctggt 2496

Qy     829 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 888
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2497 tattattcttggcgaaattcaaaggatggctcctggttcatccagtcgctttgtgccatg 2556

Qy     889 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2557 ctgaaacagtatgccgacaagcttgaatttatgcacattcttacccgggttaaccgaaag 2616

Qy     949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1008
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2617 gtggcaacagaatttgagtccttttcctttgacgctacttttcatgcaaagaaacagatt 2676

Qy    1009 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1053
           |||||||||||||||||||||||||||||||||||||||||||||
Db    2677 ccatgtattgtttccatgctcacaaaagaactctatttttatcac 2721
```

Search completed: August 16, 2002, 00:05:34
Job time: 6240 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2002, 23:59:25 ; Search time 74.75 Seconds
(without alignments)
3808.552 Million cell updates/sec

Title: US-09-613-508-3
Perfect score: 1159
Sequence: 1 GCACGAGCGGATGGGTGCTA.....GGGGACCTACTCTCATGCTG 1159

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1159	100.0	1159	3	US-08-462-969B-3	Sequence 3, Appli
2	1133.4	97.8	1158	3	US-08-591-605-1	Sequence 1, Appli
3	1118	96.5	2635	4	US-09-484-617-3	Sequence 3, Appli
4	923.2	79.7	1001	3	US-08-964-308-5	Sequence 5, Appli
5	923.2	79.7	1001	3	US-08-964-313-5	Sequence 5, Appli
6	923.2	79.7	1001	4	US-09-069-138-5	Sequence 5, Appli
7	921.6	79.5	1001	3	US-08-964-308-9	Sequence 9, Appli
8	921.6	79.5	1001	3	US-08-964-313-9	Sequence 9, Appli
9	921.6	79.5	1001	4	US-09-069-138-9	Sequence 9, Appli
10	642.4	55.4	1177	2	US-08-890-542A-1	Sequence 1, Appli
11	642.4	55.4	1177	2	US-08-890-542A-3	Sequence 3, Appli
12	631.8	54.5	1297	4	US-09-484-617-10	Sequence 10, Appl
13	229	19.8	1371	3	US-08-462-969B-1	Sequence 1, Appli
14	131.6	11.4	2658	2	US-08-773-608A-1	Sequence 1, Appli
15	131.2	11.3	1545	1	US-08-446-925-4	Sequence 4, Appli
16	131.2	11.3	1545	2	US-09-146-331-4	Sequence 4, Appli
17	131.2	11.3	1545	2	US-08-896-885-4	Sequence 4, Appli
18	131.2	11.3	1545	4	US-09-375-256-4	Sequence 4, Appli
19	126.4	10.9	1183	3	US-08-522-813-1	Sequence 1, Appli
20	98.2	8.5	1313	1	US-08-446-925-6	Sequence 6, Appli
21	98.2	8.5	1313	2	US-09-146-331-6	Sequence 6, Appli
22	98.2	8.5	1313	2	US-08-896-885-6	Sequence 6, Appli
23	98.2	8.5	1313	4	US-09-375-256-6	Sequence 6, Appli
24	55.2	4.8	2619	5	PCT-US96-10521-17	Sequence 17, Appl
25	55.2	4.8	2887	5	PCT-US96-10521-14	Sequence 14, Appl
26	53.6	4.6	870	4	US-09-360-017-2	Sequence 2, Appli
27	53.6	4.6	1286	2	US-08-807-200-3	Sequence 3, Appli
28	53.6	4.6	1286	4	US-09-001-777-3	Sequence 3, Appli
29	53.6	4.6	1330	2	US-08-807-200-1	Sequence 1, Appli
30	53.6	4.6	1330	4	US-09-001-777-1	Sequence 1, Appli
31	53.6	4.6	1355	2	US-08-618-408B-3	Sequence 3, Appli
32	53.6	4.6	1443	5	PCT-US96-10521-33	Sequence 33, Appl
33	53.6	4.6	1883	1	US-08-665-220-3	Sequence 3, Appli
34	53.6	4.6	1883	4	US-09-487-445-3	Sequence 3, Appli
35	53.6	4.6	1883	4	US-09-291-692-3	Sequence 3, Appli
36	53.6	4.6	2649	2	US-08-718-964-1	Sequence 1, Appli
37	53.6	4.6	2649	2	US-09-059-964A-1	Sequence 1, Appli
38	53.6	4.6	2649	2	US-08-842-341-1	Sequence 1, Appli
c 39	53.2	4.6	7218	1	US-08-232-463-14	Sequence 14, Appl
40	52	4.5	1440	2	US-08-807-200-11	Sequence 11, Appl
41	52	4.5	1440	3	US-08-852-782-2	Sequence 2, Appli
42	52	4.5	1440	4	US-09-001-777-11	Sequence 11, Appl
43	52	4.5	1644	3	US-08-852-782-1	Sequence 1, Appli
44	48.6	4.2	1323	5	PCT-US96-10521-15	Sequence 15, Appl
45	46.6	4.0	1301	5	PCT-US96-10521-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-08-462-969B-3
; Sequence 3, Application US/08462969B
; Patent No. 6087150
; GENERAL INFORMATION:
; APPLICANT: He, Wei-Wu et al.
; TITLE OF INVENTION: Interleukin-1 Beta Converting Enzyme
; TITLE OF INVENTION: Like Apoptosis Protease 3 and 4
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Ave.
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,969B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/334,251
; FILING DATE: 11-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF140P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1159 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-462-969B-3

Query Match 100.0%; Score 1159; DB 3; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60

```
Db      1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60

Qy     61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120
Db     61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120

Qy    121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180
Db    121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180

Qy    181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240
Db    181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240

Qy    241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300
Db    241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300

Qy    301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360
Db    301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360

Qy    361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420
Db    361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420

Qy    421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480
Db    421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480

Qy    481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540
Db    481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540

Qy    541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600
Db    541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600

Qy    601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660
Db    601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660

Qy    661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720
Db    661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720

Qy    721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780
Db    721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780

Qy    781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840
Db    781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840

Qy    841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900
Db    841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900

Qy    901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960
Db    901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960

Qy    961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020
Db    961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020

Qy   1021 TCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080
Db   1021 TCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080

Qy   1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTG 1140
```

```
Db   1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTG 1140

Qy   1141 GGGACCTACTCTCATGCTG 1159
Db   1141 GGGACCTACTCTCATGCTG 1159


RESULT   2
US-08-591-605-1
; Sequence 1, Application US/08591605
; Patent No. 6060238
;  GENERAL INFORMATION:
;    APPLICANT:  Dixit, Vishva M.
;    TITLE OF INVENTION:  METHOD AND COMPOSITION FOR REGULATING
;    TITLE OF INVENTION:  APOPTOSIS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 PAGE MILL ROAD
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/591,605
;      FILING DATE:  09-FEB-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KONSKI, ANTOINETTE F.
;      REGISTRATION NUMBER:  34,202
;      REFERENCE/DOCKET NUMBER:  20344-21036.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  706141 MRSNFOERS SFO
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1158 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  224..1054
;      OTHER INFORMATION:  /product= "Yama peptide sequence"
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  704..718
;      OTHER INFORMATION:  /note= "region coding for
;      OTHER INFORMATION:  pentapeptide, wherein cysteine is catalytic"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "atg")
;      OTHER INFORMATION:  /note= "Mutation of catalytic cysteine to methionine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gcc")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gcg")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gct")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
```

```
;      LOCATION:  replace(710..712, "gca")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
US-08-591-605-1


  Query Match             97.8%;  Score 1133.4;  DB 3;  Length 1158;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1156;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy       1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 61

Qy      61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 121

Qy     121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 181

Qy     181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 241

Qy     241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 301

Qy     301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 361

Qy     361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 421

Qy     421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     422 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 481

Qy     481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     482 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 541

Qy     541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     542 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 601

Qy     601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     602 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 661

Qy     661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     662 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 721

Qy     721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     722 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 781

Qy     781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     782 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 841

Qy     841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     842 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 901

Qy     901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     902 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 961

Qy     961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     962 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1021

Qy    1021 TCCATGCTCACAAAAGAACTCTATTTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1022 TCCATGCTCACAAAAGAACTCTATTTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1081

Qy    1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTG 1140
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    1082 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTT-GGTACTGTATTTCCCTCTCATTT 1140

Qy    1141 GGGACCTACTCTCATGCTG 1159
           ||| ||||||||||||||||
Db    1141 GGG-CCTACTCTCATGCTG 1158


RESULT   3
US-09-484-617-3
; Sequence 3, Application US/09484617
; Patent No. 6303374
; GENERAL INFORMATION:
;  APPLICANT: Hong Zhang
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF CASPASE 3 EXPRESSION
;  FILE REFERENCE: RTS-0103
;  CURRENT APPLICATION NUMBER: US/09/484,617
;  CURRENT FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 176
; SEQ ID NO 3
;   LENGTH: 2635
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (225)...(1058)
US-09-484-617-3


  Query Match             96.5%;  Score 1118;  DB 4;  Length 2635;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1143;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    2;

Qy      10 GATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCT 69
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12 gaggggtgctattgtgaggcggttgtagaagagtttcgtgagtgctcgcagctcatacct 71

Qy      70 GTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGG 129
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72 gtggctgtgtatccgtggccacagctggttggcgtcgccttgaaatcccaggccgtgagg 131

Qy     130 AGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTG 189
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     132 agttagcgagccctgctcacactcggcgctctggttttcggtgggtgtgccctgcacctg 191

Qy     190 CCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGAT 249
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     192 cctcttcccccattctcattaataaaggtatccatggagaacactgaaaactcagtggat 251

Qy     250 TCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCT 309
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     252 tcaaaatccattaaaaatttggaaccaaagatcatacatggaagcgaatcaatggactct 311

Qy     310 GGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATA 369
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     312 ggaatatccctggacaacagttataaaatggattatcctgagatgggtttatgtataata 371

Qy     370 ATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTC 429
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     372 attaataataagaattttcataaaagcactggaatgacatctcggtctggtacagatgtc 431

Qy     430 GATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAAT 489
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    432 gatgcagcaaacctcagggaaacattcagaaacttgaaatatgaagtcaggaataaaaat 491

Qy    490 GATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGC 549
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 gatcttacacgtgaagaaattgtggaattgatgcgtgatgtttctaaagaagatcacagc 551

Qy    550 AAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGA 609
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    552 aaaaggagcagttttgtttgtgtgcttctgagccatggtgaagaaggaataatttttgga 611

Qy    610 ACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGA 669
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    612 acaaatggacctgttgacctgaaaaaaataacaaacttttcagaggggatcgttgtaga 671

Qy    670 AGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGAC 729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    672 agtctaactggaaaacccaaacttttcattattcaggcctgccgtggtacagaactggac 731

Qy    730 TGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTG 789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    732 tgtggcattgagacagacagtggtgttgatgatgacatggcgtgtcataaaatacCagtg 791

Qy    790 GAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCA 849
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    792 gatgccgacttcttgtatgcatactccacagcacctggttattattcttggcgaaattca 851

Qy    850 AAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAG 909
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    852 aaggatggctcctggttcatccagtcgctttgtgccatgctgaaacagtatgccgacaag 911

Qy    910 CTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCC 969
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    912 cttgaatttatgcacattcttacccgggttaaccgaaaggtggcaacagaatttgagtcc 971

Qy    970 TTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTC 1029
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    972 ttttcctttgacgctacttttcatgcaaagaaacagattccatgtattgtttccatgctc 1031

Qy   1030 ACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTG 1089
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1032 acaaaagaactctatttttatcactaaagaaatggttggttggtggttttttttagtttg 1091

Qy   1090 TATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTAC 1149
          |||||||||||||||||||||||| |||||||||||||||||||||   |||||||
Db   1092 tatgccaagtgagaagatggtatattt-ggtactgtatttccctctcatt-ttgacctac 1149

Qy   1150 TCTCATGCTG 1159
          ||||||||||
Db   1150 tctcatgctg 1159


RESULT   4
US-08-964-308-5
; Sequence 5, Application US/08964308
; Patent No. 6066715
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBERT J. NORTH - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,308
;      FILING DATE:  04-NOV-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  NORTH, ROBERT J
;      REGISTRATION NUMBER:  27,366
;      REFERENCE/DOCKET NUMBER:  19840 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-7262
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-308-5


  Query Match           79.7%;  Score 923.2;  DB 3;  Length 1001;
  Best Local Similarity  99.5%;  Pred. No. 8.2e-273;
  Matches  947;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    2;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    529 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy     928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy     988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy    1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy    1108 GGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTACTCTCATGCTG 1159
           |||||||||  ||||||||||||||||||||  ||||||||||||||||||
Db     949 GGTATATTT-GGTACTGTATTTCCCTCTCATT-TTGACCTACTCTCATGCTG 998
```

```
RESULT   5
US-08-964-313-5
; Sequence 5, Application US/08964313
; Patent No. 6114132
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  GRESSER, MICHAEL
;    APPLICANT:  KENNEDY, BRIAN
;    APPLICANT:  NICHOLSON, DONALD
;    APPLICANT:  RAMACHANDRAN, CHIDAMBARAN
;    APPLICANT:  SKOREY, KATHRYN
;    APPLICANT:  FORD-HUTCHINSON, ANTHONY
;    TITLE OF INVENTION:  PHOSPHATASE BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,313
;      FILING DATE:  04-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/030,408
;      FILING DATE:  04-NOV-1996
;      APPLICATION NUMBER:  PCT/CA97/00825
;      FILING DATE:  03 -NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19824Y
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-313-5
```

```
  Query Match           79.7%;  Score 923.2;  DB 3;  Length 1001;
  Best Local Similarity  99.5%;  Pred. No. 8.2e-273;
  Matches  947;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    2;

Qy     208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy     268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy     328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy     388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy     448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy     508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy     568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy     628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy     688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     529 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy     748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy     808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy     868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy     928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy     988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy    1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy    1108 GGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTACTCTCATGCTG 1159
           |||||||||  ||||||||||||||||||||  ||||||||||||||||||
Db     949 GGTATATTT-GGTACTGTATTTCCCTCTCATT-TTGACCTACTCTCATGCTG 998
```

```
RESULT   6
US-09-069-138-5
; Sequence 5, Application US/09069138
; Patent No. 6348572
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  DUFRESNE, CLAUDE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  LEBLANC, YVES
;    APPLICANT:  ROY, PATRICK
;    APPLICANT:  YOUNG, ROBERT N.
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/069,138
;      FILING DATE:  29-APR-1998
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19840YIA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-069-138-5


  Query Match             79.7%;  Score 923.2;  DB 4;  Length 1001;
  Best Local Similarity   99.5%;  Pred. No. 8.2e-273;
  Matches  947;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    2;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    529 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy    988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy   1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy   1108 GGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTACTCTCATGCTG 1159
          ||||||||| |||||||||||||||||||||    ||||||||||||||||
Db    949 GGTATATTT-GGTACTGTATTTCCCTCTCATT-TTGACCTACTCTCATGCTG 998


RESULT   7
US-08-964-308-9
; Sequence 9, Application US/08964308
; Patent No. 6066715
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBERT J. NORTH - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,308
;      FILING DATE:  04-NOV-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  NORTH, ROBERT J
;      REGISTRATION NUMBER:  27,366
;      REFERENCE/DOCKET NUMBER:  19840 PCT
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  732-594-7262
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-308-9


  Query Match           79.5%;  Score 921.6;  DB 3;  Length 1001;
  Best Local Similarity  99.4%;  Pred. No. 2.5e-272;
  Matches  946;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    529 AAACTTTTCATTATTCAGGCCTCCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy    988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy   1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy   1108 GGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTACTCTCATGCTG 1159
          ||||||||| |||||||||||||||||||||||    |||||||||||||||||
Db    949 GGTATATTT-GGTACTGTATTTCCCTCTCATT-TTGACCTACTCTCATGCTG 998


RESULT   8
US-08-964-313-9
; Sequence 9, Application US/08964313
; Patent No. 6114132
; GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  GRESSER, MICHAEL
;    APPLICANT:  KENNEDY, BRIAN
;    APPLICANT:  NICHOLSON, DONALD
;    APPLICANT:  RAMACHANDRAN, CHIDAMBARAN
;    APPLICANT:  SKOREY, KATHRYN
;    APPLICANT:  FORD-HUTCHINSON, ANTHONY
;    TITLE OF INVENTION:  PHOSPHATASE BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,313
;      FILING DATE:  04-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/030,408
;      FILING DATE:  04-NOV-1996
;      APPLICATION NUMBER:  PCT/CA97/00825
;      FILING DATE:  03 -NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19824Y
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-313-9


  Query Match           79.5%;  Score 921.6;  DB 3;  Length 1001;
  Best Local Similarity  99.4%;  Pred. No. 2.5e-272;
  Matches  946;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    529 AAACTTTTCATTATTCAGGCCTCCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy    988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy   1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy   1108 GGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTACTCTCATGCTG 1159
          |||||||| |||||||||||||||||||||   |||||||||||||||||
Db    949 GGTATATTT-GGTACTGTATTTCCCTCTCATT-TTGACCTACTCTCATGCTG 998

RESULT   9
US-09-069-138-9
; Sequence 9, Application US/09069138
; Patent No. 6348572
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  DUFRESNE, CLAUDE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  LEBLANC, YVES
;    APPLICANT:  ROY, PATRICK
;    APPLICANT:  YOUNG, ROBERT N.
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/069,138
;      FILING DATE:  29-APR-1998
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19840YIA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-069-138-9


  Query Match           79.5%;  Score 921.6;  DB 4;  Length 1001;
  Best Local Similarity  99.4%;  Pred. No. 2.5e-272;
  Matches  946;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
```

```
Db     529 AAACTTTTCATTATTCAGGCCTCCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy     748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy     808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy     868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy     928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy     988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy    1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy    1108 GGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTACTCTCATGCTG 1159
           |||||||| ||||||||||||||||||||    ||||||||||||||||||
Db     949 GGTATATTT-GGTACTGTATTTCCCTCTCATT-TTGACCTACTCTCATGCTG 998


RESULT  10
US-08-890-542A-1
; Sequence 1, Application US/08890542A
; Patent No. 5840509
;  GENERAL INFORMATION:
;    APPLICANT:  Ni, Binhui
;    APPLICANT:  Paul, Steven M
;    APPLICANT:  Wu, Xin
;    TITLE OF INVENTION:  PROTEASE AND RELATED NUCLEIC ACID
;    TITLE OF INVENTION:  COMPOUNDS
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  United States of America
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/890,542A
;      FILING DATE:  09-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gaylo, Paul J.
;      REGISTRATION NUMBER:  36,808
;      REFERENCE/DOCKET NUMBER:  X-10704
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (317) 276-0756
;      TELEFAX:  (317) 276-3861
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1177 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  76..906
US-08-890-542A-1


  Query Match             55.4%;  Score 642.4;  DB 2;  Length 1177;
  Best Local Similarity   83.8%;  Pred. No. 8e-187;
  Matches  727;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy     210 AATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTT 269
           || ||| ||  |||||||| ||||  |||| ||| ||||||||||||||||||||| |||||
Db      63 AAGAAAAGTGACCATGGACAACAACGAAACCTCCGTGGATTCAAAATCCATTAATAATTT 122

Qy     270 GGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAG 329
           ||| ||||||  || |||||||| | || ||||||||||||||||  ||||||| |||
Db     123 TGAAACAAAGACTATCCATGGAAGCAAGTCGATGGACTCTGGAATATATCTGGACAGCAG 182

Qy     330 TTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCA 389
           ||| |||||||||| ||||| ||||| || ||||||||||||||||||||||| || ||
Db     183 TTACAAAATGGATTACCCTGAAATGGGCTTGTGTATAATAATTAATAATAAGAACTTCCA 242

Qy     390 TAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGA 449
           |||||||||||||||| || ||||  |||||| ||||||||||||||| ||||||| ||
Db     243 TAAAAGCACTGGAATGTCAGCTCGCAATGGTACCGATGTCGATGCAGCTAACCTCAGAGA 302

Qy     450 AACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAAT 509
           |||||||    | ||||||| ||||||||||||||||||||| ||||| ||||||||||
Db     303 GACATTCATGGCCCTGAAATACGAAGTCAGGAATAAAAATGACCTTACTCGTGAAGAAAT 362

Qy     510 TGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTG 569
           | |||||||||||  | |||||||| ||||||||||||||||||||||||||||| ||
Db     363 TATGGAATTGATGGATAGTGTTTCTAAGGAAGATCACAGCAAAAGGAGCAGTTTTGTGTG 422

Qy     570 TGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCT 629
           |||| |||| || ||||| || ||||| ||||||||||||| || |||||||| |||||
Db     423 TGTGATTCTAAGTCATGGAGATGAAGGAGTAATTTTTGGAACGAACGGACCTGTGGACCT 482

Qy     630 GAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAA 689
           ||||||| |||| |  || |||||||| ||    ||  | ||||| |||||||| || ||
Db     483 GAAAAAACTAACTAGTTTCTTCAGAGGCGACTACTGCCGGAGTCTGACTGGAAAGCCGAA 542

Qy     690 ACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAG 749
           ||| ||||| |||||||||||||| |||||||| |||||||| || ||||||||||||||
Db     543 ACTCTTCATCATTCAGGCCTGCCGAGGTACAGAGCTGGACTGCGGTATTGAGACAGACAG 602

Qy     750 TGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGC 809
           |||   ||| ||||| ||||| || || || |||||||||| ||||||||||| ||||||
Db     603 TGGAACTGACGATGATATGGCATGCCAGAAGATACCAGTGGGGGCCGACTTCCTGTATGC 662

Qy     810 ATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCAT 869
           ||||| || ||||| ||||| ||||| ||| ||||||||| ||| || || ||||||||
Db     663 TTACTCTACCGCACCCGGTTACTATTCCTGGAGAAATTCAAGGGACGGGTCATGGTTCAT 722

Qy     870 CCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCT 929
           |||||| |||| |||||||||||||| ||| ||  ||||||| ||||| |||||||| ||
Db     723 CCAGTCACTTTGCGCCATGCTGAAACTGTACGCGCACAAGCTGGAATTCATGCACATCCT 782

Qy     930 TACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTT 989
           || || |||||||| |||||||| | |||||||||||||| ||| | |||| |||||
Db     783 CACTCGTGTTAACCGGAAGGTGGCCATGGAATTTGAGTCCTTCTCCCTGGACGCCACTTT 842

Qy     990 TCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTA 1049
           || |||||||||||||| || |||||||| || ||||||||||||||||| || ||||||
Db     843 CCACGCAAAGAAACAGATCCCGTGTATTGTGTCAATGCTCACAAAAGAACTGTACTTTTA 902

Qy    1050 TCACTAAAGAAATGGTTGGTTGGTGGTT 1077
           ||||||||| |||| ||| |  || | |
Db     903 TCACTAAAGGAATGACTGGGAGTGGGGT 930


RESULT  11
US-08-890-542A-3
; Sequence 3, Application US/08890542A
```

```
; Patent No. 5840509
;  GENERAL INFORMATION:
;    APPLICANT:  Ni, Binhui
;    APPLICANT:  Paul, Steven M
;    APPLICANT:  Wu, Xin
;    TITLE OF INVENTION:  PROTEASE AND RELATED NUCLEIC ACID
;    TITLE OF INVENTION:  COMPOUNDS
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  United States of America
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/890,542A
;      FILING DATE:  09-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gaylo, Paul J.
;      REGISTRATION NUMBER:  36,808
;      REFERENCE/DOCKET NUMBER:  X-10704
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (317) 276-0756
;      TELEFAX:  (317) 276-3861
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1177 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  mRNA
US-08-890-542A-3


  Query Match           55.4%;  Score 642.4;  DB 2;  Length 1177;
  Best Local Similarity   62.0%;  Pred. No. 8e-187;
  Matches  538;  Conservative  189;  Mismatches  141;  Indels    0;  Gaps    0;

Qy    210 AATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTT 269
          || ||| |:  |||:||| |||| |||| |:| |:|||::|||||:|||::|| ||::: 
Db     63 AAGAAAAGUGACCAUGGACAACAACGAAACCUCCGUGGAUUCAAAAUCCAUUAAUAAUUU 122

Qy    270 GGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAG 329
           ||| ||||||  |: ||:|||||| | :| |:||||:|:|||||:|:  |:||||| ||| 
Db    123 UGAAACAAAGACUAUCCAUGGAAGCAAGUCGAUGGACUCUGGAAUAUAUCUGGACAGCAG 182

Qy    330 TTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCA 389
          ::| ||||:|||::| ||:|| |:|||  :: :|:|:||:||::||:||:||||| :: || 
Db    183 UUACAAAAUGGAUUACCCUGAAAUGGGCUUGUGUAUAAUAAUUAAUAAUAAGAACUUCCA 242

Qy    390 TAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGA 449
          :|||||||||:||||:| || |:||   :||:|| ||:|:|||:||||| ||||:||| || 
Db    243 UAAAAGCACUGGAAUGUCAGCUCGCAAUGGUACCGAUGUCGAUGCAGCUAACCUCAGAGA 302

Qy    450 AACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAAT 509
           |||::||    | :||||:| ||||:|||||:|||||:|| |:::|| ||:|||||: 
Db    303 GACAUUCAUGGCCCUGAAAUACGAAGUCAGGAAUAAAAAUGACCUUACUCGUGAAGAAAU 362

Qy    510 TGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTG 569
          : :|||||::||:|  :  :|:::|:|| |||||:|||||||||||||||:::|: :| 
Db    363 UAUGGAAUUGAUGGAUAGUGUUUCUAAGGAAGAUCACAGCAAAAGGAGCAGUUUUGUGUG 422

Qy    570 TGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCT 629
          :|:| ::|: || ||:|| || |||||| :||:::::|||||| || |||||:|: |||||
Db    423 UGUGAUUCUAAGUCAUGGAGAUGAAGGAGUAAUUUUUGGAACGAACGGACCUGUGGACCU 482

Qy    630 GAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAA 689
          ||||||| :||| |  :: ::|||||| ||    :|  | ||:|: ||:|||| || || 
Db    483 GAAAAAACUAACUAGUUUCUUCAGAGGCGACUACUGCCGGAGUCUGACUGGAAAGCCGAA 542

Qy    690 ACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAG 749
          ||: ::||: |::|||||:|||| ||:||||| |:||||:| || |::|||||||||| 
Db    543 ACUCUUCAUCAUUCAGGCCUGCCGAGGUACAGAGCUGGACUGCGGUAUUGAGACAGACAG 602

Qy    750 TGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGC 809
          :||   :|| ||:|| |:||| :| || || |:|||||:|| |||||||::| :|:|:|| 
Db    603 UGGAACUGACGAUGAUAUGGCAUGCCAGAAGAUACCAGUGGGGGCCGACUUCCUGUAUGC 662

Qy    810 ATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCAT 869
           :||:| || ||||| ||::| :|::| :|| ||||::||| ||| || :| :|||::||: 
Db    663 UUACUCUACCGCACCCGGUUACUAUUCCUGGAGAAAUUCAAGGGACGGGUCAUGGUUCAU 722

Qy    870 CCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCT 929
          ||||:| |:::| ||||:||:||||| |:| ||  |||||: |||:: |:|||||: |: 
Db    723 CCAGUCACUUUGCGCCAUGCUGAAACUGUACGCGCACAAGCUGGAAUUCAUGCACAUCCU 782

Qy    930 TACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTT 989
           || || |::|||| ||||:||| |  |||:::|||:||:: :|| : ||||| ||::: 
Db    783 CACUCGUGUUAACCGGAAGGUGGCCAUGGAAUUUGAGUCCUUCUCCCUGGACGCCACUUU 842

Qy    990 TCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTA 1049
          || ||||||||||||: || :|:|::|: :| |:||:|||||||||||: :| ::::| 
Db    843 CCACGCAAAGAAACAGAUCCCGUGUAUUGUGUCAAUGCUCACAAAAGAACUGUACUUUUA 902

Qy   1050 TCACTAAAGAAATGGTTGGTTGGTGGTT 1077
          :|||:|||| ||:|  :||  |  || : 
Db    903 UCACUAAAGGAAUGACUGGGAGUGGGGU 930


RESULT  12
US-09-484-617-10
; Sequence 10, Application US/09484617
; Patent No. 6303374
; GENERAL INFORMATION:
;  APPLICANT: Hong Zhang
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF CASPASE 3 EXPRESSION
;  FILE REFERENCE: RTS-0103
;  CURRENT APPLICATION NUMBER: US/09/484,617
;  CURRENT FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 176
; SEQ ID NO 10
;   LENGTH: 1297
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (68)...(901)
US-09-484-617-10


  Query Match           54.5%;  Score 631.8;  DB 4;  Length 1297;
  Best Local Similarity   83.0%;  Pred. No. 1.5e-183;
  Matches  720;  Conservative    0;  Mismatches  147;  Indels    0;  Gaps    0;

Qy    209 TAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATT 268
          ||| ||| ||  |||||||||||   ||| ||||||||||||||||||||||||| |||| 
Db     54 taagaaaagtgaccatggagaacaacaaaacctcagtggattcaaaatccattaataatt 113

Qy    269 TGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 328
          | |||  ||||| |||||||||| ||| | ||| |||||||||| || |  |||||||  | 
Db    114 ttgaagtaaagaccatacatgggagcaagtcagtggactctgggatctatctggacagta 173

Qy    329 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTC 388
          |||| |||||||||||||||| |||||  |||| |||||||||||||||||||| || | 
Db    174 gttacaaaatggattatcctgaaatgggcatatgcataataattaataataagaacttcc 233

Qy    389 ATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGG 448
          |||| |||||||||||| ||||||| |||||||| ||||| || ||||| |||||||| | 
```

```
Db    234 ataagagcactggaatgtcatctcgctctggtacggatgtggacgcagccaacctcagag 293

Qy    449 AAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAA 508
          | |||||||    | |||||||  |||||||||||||||||||||||| ||||||||
Db    294 agacattcatgggcctgaaataccaagtcaggaataaaaatgatcttactcgtgaagaca 353

Qy    509 TTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTT 568
          || ||||||| |||  |  |||||||| |||||||| |||||||||||||| ||||| |
Db    354 ttttggaattaatggatagtgtttctaaggaagatcatagcaaaaggagcagctttgtgt 413

Qy    569 GTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 628
          ||||| |||| |||||||||| |||||  | |||| ||| |||||||| |||||||| |
Db    414 gtgtgattctaagccatggtgatgaaggggtcatttatgggacaaatgggcctgttgaac 473

Qy    629 TGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCA 688
          ||||||| | || | ||| |||||||| ||    ||  | ||||| |||||||| || |
Db    474 tgaaaaagttgactagcttcttcagaggcgactactgccggagtctgactggaaagccga 533

Qy    689 AACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACA 748
          |||| ||||| |||||||||||||| ||||| || ||||||||||||||||||||||||
Db    534 aactcttcatcattcaggcctgccggggtacggagctggactgtggcattgagacagaca 593

Qy    749 GTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATG 808
          ||||  |||||| || ||||| || || || ||||| |||||||| |||||| ||||||
Db    594 gtgggactgatgaggagatggcttgccagaagatacggtggaggctgacttcctgtatg 653

Qy    809 CATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCA 868
          | ||||| |||||||||||||| ||||| ||| |||||||||||| || || |||||||
Db    654 cttactctacagcacctggttactattcctggagaaattcaaaggacgggtcgtggttca 713

Qy    869 TCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 928
          ||||||| |||||  |||||||||| | ||| ||  |||||||||||||||||||||
Db    714 tccagtccctttgcagcatgctgaagctgtacgcgcacaagctagaatttatgcacattc 773

Qy    929 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTT 988
          | || || |||||| | |||||||||| ||||| |||||||| ||| | ||| | ||||
Db    774 tcactcgcgttaacaggaaggtggcaacggaattcgagtccttctccctggactccactt 833

Qy    989 TTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTT 1048
          | || ||||||||||||||| || |||||||| |||||||||| |||||||| || ||||
Db    834 tccacgcaaagaaacagatcccgtgtattgtgtccatgctcacgaaagaactgtactttt 893

Qy   1049 ATCACTAAAGAAATGGTTGGTTGGTGG 1075
          ||||||| || |||| ||||  |  ||
Db    894 atcactagaggaatgattgggggtggg 920


RESULT 13
US-08-462-969B-1
; Sequence 1, Application US/08462969B
; Patent No. 6087150
;  GENERAL INFORMATION:
;    APPLICANT:  He, Wei-Wu et al.
;    TITLE OF INVENTION:  Interleukin-1 Beta Converting Enzyme
;    TITLE OF INVENTION:  Like Apoptosis Protease 3 and 4
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Ave.
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,969B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/334,251
;      FILING DATE:  11-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF140P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1371 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-462-969B-1


  Query Match             19.8%;  Score 229;  DB 3;  Length 1371;
  Best Local Similarity   58.6%;  Pred. No. 3e-60;
  Matches  439;  Conservative    0;  Mismatches  300;  Indels   10;  Gaps    2;

Qy    331 TATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCAT 390
          || || ||| ||| |    || |||   ||| || ||||| || || ||||| ||| ||
Db    517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576

Qy    391 AAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAA 450
          |||   || || |||    |||   || |||||   ||||| |    |||    |
Db    577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636

Qy    451 ACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATT 510
             ||| ||| | ||   | ||| || |       | |||||     |  |||   | ||
Db    637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696

Qy    511 GTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGT 570
             ||  || |      | | |||| |||| || || |  ||       | | || |  ||
Db    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756

Qy    571 GTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTG 630
             | ||  | |||||||| ||||||    |||||| ||| | | ||||           |
Db    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816

Qy    631 AAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAA 690
          ||  |  | |||  |   || || |||||| | || | ||  ||     |  |||||||||
Db    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876

Qy    691 CTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGT 750
          || ||| | |||||||| ||||| || || || || ||   |||||  || | |||
Db    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936

Qy    751 GG---------TGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTC 801
          ||         ||      |||| |   |  | || || |||||||| || ||||||
Db    937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996

Qy    802 TTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCC 861
          | |  || || ||||| |  || || ||||| || ||| | |   ||      ||||||
Db    997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056

Qy    862 TGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATG 921
          |||||  | ||  | || ||  |||| ||| |  || | |   |  | || ||| | |||
Db   1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116

Qy    922 CACATTCT-TACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGA 980
          || || ||  ||| |||| ||     | || || |  | ||||||||    ||   |||
Db   1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176

Qy    981 CGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACT 1040
          | |    || ||||  ||||| ||||| || ||| | || |||||||||| || |||||
Db   1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236

Qy   1041 CTATTTTTATCACTAAAGAAATGGTTGGT 1069
          ||| ||   ||| ||   | ||    |||
```

```
Db   1237 CTACTTCAGTCAATAGCCATATCAGGGGT 1265


RESULT  14
US-08-773-608A-1
; Sequence 1, Application US/08773608A
; Patent No. 5858778
;  GENERAL INFORMATION:
;    APPLICANT:  Alnemri, Emad S.
;    APPLICANT:  Fernandes-Alnemri, Teresa
;    APPLICANT:  Litwack, Gerald
;    TITLE OF INVENTION:  Sf Caspase-1, Compositions and Methods
;    TITLE OF INVENTION:  for Making and Methods of Using the Same
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5858778ris
;      STREET:  One Liberty Place 46th floor
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  US
;      ZIP:  19355
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/773,608A
;      FILING DATE:  27-DEC-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeLuca, Mark
;      REGISTRATION NUMBER:  33,229
;      REFERENCE/DOCKET NUMBER:  TJU-2087
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2658 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  140..1039
US-08-773-608A-1


  Query Match             11.4%;  Score 131.6;  DB 2;  Length 2658;
  Best Local Similarity   52.1%;  Pred. No. 2.9e-30;
  Matches  371;  Conservative    0;  Mismatches  329;  Indels   12;  Gaps    3;

Qy    330 TTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCA 389
          ||| || ||| |  ||    |   ||| |    || ||  | ||  |  || ||||  |
Db    295 TTACAACATGAACCATAAACATCGTGGTATGGCCATTATTTTCAACCACGAGCATTTCGA 354

Qy    390 TAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGA 449
          |    ||   |  ||| || ||  |||| ||  |||| ||    |  |||||     |
Db    355 CATTCACA---GCCTGAAGTCCCGTACTGGCACGAATGTAGACAGCGACAACCTTTCCAA 411

Qy    450 AACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAAT 509
          |    | | || ||||  ||   |||| | |        | |  | | |   ||||||||
Db    412 AGTTCTTAAAACCTTGGGATTCAAAGTTACGGTGTTCCCTAACTTAAAATCAGAAGAAAT 471

Qy    510 TGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTG 569
              |||| || |   |   | |  |    || ||     |       | ||||  ||
Db    472 CAATAAATTTATCCAGCAGACTGCAGAGATGGACCATTCAGACGCTGACTGTTTACTTGT 531

Qy    570 TGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCT 629
          ||   || | | |||||| ||   ||||||  | | ||  | | ||   | |    | |
Db    532 TGCCGTTTTAACCCATGGAGAGCTAGGAATGCTGTATGCCAAAGATACTCATTACAAGCC 591

Qy    630 GAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAA 689
          | ||  |     |||  || |  |  ||    ||    |  ||  ||||||||||  ||
Db    592 AGACAACCTGTGGTACTACTTTACTGCTGACAAATGCCCCACACTGGCTGGAAAACCTAA 651

Qy    690 ACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAG 749
          || ||| ||||||||||| ||||  |||     | ||||  |||| |||        | |
Db    652 GCTCTTCTTTATTCAGGCTTGCCAAGGTGACAGATTGGATGGTGGTATTACTCTGAGCCG 711

Qy    750 TGGTGTTGATGATGACATGGCGTGT------CATAAAATACCAGTGGAGGCCGACTTCTT 803
            ||     |||| |    | | |       |||  || || ||  | || || |||||
Db    712 CACTGAGACAGATGGCTCACCTTCTACTTCGTATAGGATTCCCGTACATGCTGATTTCTT 771

Qy    804 GTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTG 863
          |  |||||  || || | |||||| || | |||||| | ||  | |     || || ||
Db    772 GATTGCATTTTCAACTGTACCTGGATACTTTTCTTGGAGGAACACTACCAGAGGTTCATG 831

Qy    864 GTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCA 923
          |||| ||  ||  |  | ||||    |||     |    |       ||   ||  || ||
Db    832 GTTTATGCAAGCTTTGTGTGAAGAGCTACGCTATGCAGGAACAGAGAGGGACATTCTGAC 891

Qy    924 CATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGAC-- 981
          | || ||    ||    |  ||||| ||   ||| ||||||||   | ||  ||||
Db    892 ATTGCTGACTTTTGTGTGTCAGAAGGTTGCGCTAGACTTTGAGTCTAATGCCCCTGACTC 951

Qy    982 -GCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACA 1032
           || |   | |||  | |||||||  |||| || ||      ||||||||||
Db    952 AGCGATGATGCATCAACAGAAACAAGTTCCCTGCATCACCAGCATGCTCACA 1003


RESULT  15
US-08-446-925-4
; Sequence 4, Application US/08446925
; Patent No. 5672500
;  GENERAL INFORMATION:
;    APPLICANT:  Litwack, Gerald
;    APPLICANT:  Alnemri, Emad S.
;    APPLICANT:  Fernandez-Alnemri, Teresa
;    TITLE OF INVENTION:  Mch2, AN APOPTOTIC CYSTEINE
;    TITLE OF INVENTION:  PROTEASE,
;    TITLE OF INVENTION:  AND COMPOSITIONS FOR MAKING AND
;    TITLE OF INVENTION:  METHODS
;    TITLE OF INVENTION:  OF USING THE SAME
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewicz &
;      ADDRESSEE:  No. 5672500ris
;      STREET:  One Liberty Place, 46th floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/446,925
;      FILING DATE:
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeLuca, Mark
;      REGISTRATION NUMBER:  33,229
;      REFERENCE/DOCKET NUMBER:  TJU-1508
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1545 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
```

```
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  79..957
US-08-446-925-4


  Query Match            11.3%;  Score 131.2;  DB 1;  Length 1545;
  Best Local Similarity  50.5%;  Pred. No. 2.8e-30;
  Matches  405;  Conservative    0;  Mismatches  358;  Indels   39;  Gaps    2;

Qy     319 CTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 378
           | ||   | |  || ||||||||  |      ||  ||  |   | ||||  | ||| ||
Db     175 CCGGCAGAAAAGTACAAAATGGACCACAGGAGGAGAGGAATTGCTTTAATCTTCAATCAT 234

Qy     379 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCA 438
            |||  ||   \     \     \ || ||    ||    \ ||    |  |||  ||
Db     235 GAGAGGTTCTTTTGGCACTTAACACTGCCAGAAAGGCGGCGCACCTGCGCAGATAGAGAC 294

Qy     439 AACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACA 498
           || || |     |  ||   | |  |   || ||||||  |         ||||||||| |
Db     295 AATCTTACCCGCAGGTTTTCAGATCTAGGATTTGAAGTGAAATGCTTTAATGATCTTAAA 354

Qy     499 CGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGC 558
           ||||||  |  ||     || | ||| || || |  |     |||     |
Db     355 GCAGAAGAACTACTGCTCAAAATTCATGAGGTGTCAACTGTTAGCCACGCAGATGCCGAT 414

Qy     559 AGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGA 618
           | ||||| |||||  | ||||||||||| ||||        |||| || |    |||
Db     415 TGCTTTGTGTGTGTCTTCCTGAGCCATGGCGAAGGCAATCACATTTATGCATATGATGCT 474

Qy     619 CCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 678
           | ||  |  | | | ||||   ||| |||| ||| ||    |||   || ||   |
Db     475 AAAATCGAAATTCAGACATTAACTGGCTTGTTCAAAGGAGACAAGTGTCACAGCCTGGTT 534

Qy     679 GGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAAC-------------- 724
           ||||||||||||  | || || || ||||| || || || |   | |
Db     535 GGAAAACCCAAGATATTTATCATCCAGGCATGTCGGGGAAACCAGCACGATGTGCCAGTC 594

Qy     725 ------TGGACTGTGGCATTGAGACAGACAGTGGTGTT------------------GAT 759
               | |  ||| | |   |  |||||||  |  |||                   |
Db     595 ATTCCTTTGGATGTAGTAGATAATCAGACAGAGAAGTTGGACACCAACATAACTGAGGTG 654

Qy     760 GATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACA 819
           ||||      |   | | |   | || |  |  || ||||||  |       |||||
Db     655 GATGCAGCCTCCGTTTACACGCTGCCTGCTGGAGCTGACTTCCTCATGTGTTACTCTGTT 714

Qy     820 GCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTT 879
           |||   || ||||||||||   ||  |  |   | | ||||| |||| ||| ||     |
Db     715 GCAGAAGGATATTATTCTCACCGGGAAACTGTGAACGGCTCATGGTACATTCAAGATTTG 774

Qy     880 TGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTT 939
           ||||  ||| ||  | | |||| |  |    | || || |   |  | || || | |||
Db     775 TGTGAGATGTTGGGAAAATATGGCTCCTCCTTAGAGTTCACAGAACTCCTCACACTGGTG 834

Qy     940 AACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAG 999
           ||| | || ||  |        | |  ||||| |   ||| | | |     || ||||
Db     835 AACAGGAAAGTTTCTCAGCGCCGAGTGGACTTTTGCAAAGACCCAAGTGCAATTGGAAAG 894

Qy    1000 AAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGA 1059
           || ||| |||| ||| |||  || |||||  || ||| | ||  |||| | ||   ||
Db     895 AAGCAGGTTCCCTGTTTTGCCTCAATGCTAACTAAAAAGCTGCATTTCTTTCCAAAATCT 954

Qy    1060 AATGGTTGGTTGGTGGTTTTTT 1081
           |||   |   | | || | | |
Db     955 AATTAATTAATAGAGGCTATCT 976


Search completed: August 15, 2002, 23:59:44
Job time: 6190 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2002, 23:02:07 ; Search time 2425.75 Seconds
(without alignments)
6448.715 Million cell updates/sec

Title: US-09-613-508-3
Perfect score: 1159
Sequence: 1 GCACGAGCGGATGGGTGCTA.....GGGGACCTACTCTCATGCTG 1159

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	814.2	70.3	913	10	BM461940	BM461940 AGENCOURT
	2	780	67.3	827	9	AL534526	AL534526 AL534526
	3	722.6	62.3	846	9	AU119346	AU119346 AU119346
	4	699.8	60.4	951	9	AU133051	AU133051 AU133051
	5	654.4	56.5	826	9	AU127381	AU127381 AU127381
	6	651.4	56.2	911	10	BE780308	BE780308 601468289
	7	618.8	53.4	751	9	AU127028	AU127028 AU127028
	8	601	51.9	624	9	AV649560	AV649560 AV649560
	9	596.2	51.4	624	9	AV649613	AV649613 AV649613
	10	579.8	50.0	946	10	BG177093	BG177093 602313329
	11	576.2	49.7	965	10	BG170404	BG170404 602322734
	12	545.6	47.1	878	10	BI080567	BI080567 602877052
	13	523.2	45.1	794	10	BI652204	BI652204 603299029
	14	521.8	45.0	819	9	AU125557	AU125557 AU125557
	15	496.4	42.8	602	10	BF966141	BF966141 602286412
	16	495.4	42.7	497	9	AA984697	AA984697 am89h06.s
	17	481.6	41.6	599	9	AW951022	AW951022 EST363092
	18	465.2	40.1	638	10	BI691090	BI691090 603309866
	19	465	40.1	653	10	BI657919	BI657919 603282147
	20	454.8	39.2	523	9	AL602306	AL602306 DKFZp686J
	21	445.4	38.4	565	9	AL602756	AL602756 DKFZp686H
	22	432.8	37.3	700	10	BG865023	BG865023 602799491
	23	426.8	36.8	1007	10	BG287359	BG287359 602382123
c	24	414	35.7	445	9	AA011446	AA011446 zi01g12.s
	25	391.2	33.8	479	9	AU127408	AU127408 AU127408
	26	388	33.5	438	9	AA011445	AA011445 zi01g12.r
	27	380.8	32.9	387	9	AV648018	AV648018 AV648018
	28	357	30.8	399	10	T10341	T10341 seq1254 b4H
	29	354.8	30.6	441	10	H29199	H29199 ym31f11.r1
	30	338	29.2	705	10	BE284569	BE284569 601087184
	31	321	27.7	435	10	BI132775	BI132775 AR035B10P
	32	317.2	27.4	802	10	BI153295	BI153295 602917153
	33	313.6	27.1	352	9	AA641875	AA641875 ns19h11.r
	34	310.6	26.8	414	10	BM149220	BM149220 TCAAP2Q11
	35	305.8	26.4	411	9	AA243814	AA243814 zr67e07.r
	36	282.4	24.4	329	9	AA298723	AA298723 EST114329
	37	282	24.3	317	9	AA297850	AA297850 EST113482
	38	268.4	23.2	430	9	AI116437	AI116437 ue98b08.y
c	39	262	22.6	378	9	BE178741	BE178741 PM4-HT060
	40	261.6	22.6	513	9	AW318387	AW318387 um96a09.y
	41	258.2	22.3	263	9	AA641754	AA641754 ns19c06.r
	42	249.8	21.6	1098	10	BG392166	BG392166 602410279
	43	249.4	21.5	388	9	AI037098	AI037098 ue19d02.y
c	44	248.2	21.4	402	9	AW542198	AW542198 C0154G07-
c	45	242.2	20.9	424	10	BF825247	BF825247 CM1-HN001

ALIGNMENTS

RESULT 1
BM461940
LOCUS BM461940 913 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6418271 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5534002 5', mRNA sequence.
ACCESSION BM461940
VERSION BM461940.1 GI:18510980
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 913)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12219 row: m column: 11
High quality sequence stop: 671.
FEATURES Location/Qualifiers
source 1. .913
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5534002"
/clone_lib="NIH_MGC_71"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb. "
BASE COUNT 285 a 159 c 203 g 265 t 1 others
ORIGIN

```
  Query Match             70.3%;  Score 814.2;  DB 10;  Length 913;
  Best Local Similarity   96.3%;  Pred. No. 1.3e-201;
  Matches  855;  Conservative    0;  Mismatches   29;  Indels    4;  Gaps    2;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     24 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 83

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     84 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 143

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    144 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 203

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    204 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 263

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    264 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 323

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    324 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 383

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    384 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 443

Qy    628 CTGAAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    444 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 503

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    504 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 563

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    564 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 623

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAGGATGGCTCCTGGTTC 867
          |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db    624 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTNCAAGGATGGCTCCTGGTTC 683

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    684 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 743

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    744 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 803

Qy    988 TTTCATGCAAAGAAACAGATTCCATGTATTG-TTTCCATGCTCACAAAAGAA---CTCTA 1043
          |||  ||| ||  ||||||||||||||||| ||||||||||| | |||| |   || |
Db    804 TTTTCTGCCAAAGAACAGATTCCATGTATTGTTTTCCATGCTCCCCAAAGGAACTCTATT 863

Qy   1044 TTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTA 1091
          ||||  |||||| |||||| | |   |||  |||| ||| ||
Db    864 TTTTATCCACTAAAAAAATGGGTTGGGCTTGGGGGCTTTTTTTTTTA 911


RESULT  2
AL534526
LOCUS       AL534526                827 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL534526 LTI_FL013_FBrn1 Homo sapiens cDNA clone CS0DF004YA06 5
            prime, mRNA sequence.
ACCESSION   AL534526
VERSION     AL534526.1  GI:12798019
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 827)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .827
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DF004YA06"
                     /clone_lib="LTI_FL013_FBrn1"
                     /dev_stage="pooled tissue from post conception fetuses (20
                     week, 24 week and 26 week)"
                     /lab_host="DH10B"
                     /note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand
                     cDNA was primed with a NotI-oligo(dT) primer. Five prime
                     end enriched, double-stranded cDNA was digested with Not I
                     and cloned into the Not I and Eco RV sites of the
                     pCMVSPORT 6 vector.  Library was constructed by Life
                     Technologies.  Contact : Feng Liang Life Technologies, a
                     division of Invitrogen 9800 Medical Center Drive Rockville
                     , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
                     fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      266 a    140 c    192 g    229 t
ORIGIN


  Query Match             67.3%;  Score 780;  DB 9;  Length 827;
  Best Local Similarity   99.9%;  Pred. No. 1e-192;
  Matches  791;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     36 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 95

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     96 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 155

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    156 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 215

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    216 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 275

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    276 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 335

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    336 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 395

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    396 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 455

Qy    628 CTGAAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    456 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 515
```

```
Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    516 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 575

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    576 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 635

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    636 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 695

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    696 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 755

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAA-TTTGAGTCCTTTTCCTTTGACGCTAC 986
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    756 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTTGAGTCCTTTTCCTTTGACGCTAC 815

Qy    987 TTTTCATGCAAA 998
          ||||||||||||
Db    816 TTTTCATGCAAA 827
```

```
RESULT   3
AU119346
LOCUS       AU119346                846 bp    mRNA    linear   EST 19-OCT-2000
DEFINITION  AU119346 HEMBA1 Homo sapiens cDNA clone HEMBA1005587 5', mRNA
            sequence.
ACCESSION   AU119346
VERSION     AU119346.1  GI:10934581
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 846)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .846
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEMBA1005587"
                     /clone_lib="HEMBA1"
                     /tissue_type="whole embryo, mainly head"
                     /dev_stage="embryo, 10 weeks"
                     /note="Vector: pME18SFL3"
BASE COUNT      266 a    150 c    200 g    226 t      4 others
ORIGIN
```

```
  Query Match           62.3%;  Score 722.6;  DB 9;  Length 846;
  Best Local Similarity  96.4%;  Pred. No. 9.9e-178;
  Matches  770;  Conservative    0;  Mismatches   23;  Indels    6;  Gaps    3;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     53 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 112

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    113 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 172

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    173 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 232

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    233 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 292

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    293 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAGTGATCTTACACGTGAAGAA 352

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    353 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 412

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    413 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 472

Qy    628 CTGAAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    473 CTGAAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 532

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    533 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 592

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    593 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTNAT 652

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          |||||||||||||||||||| |||||||||||||||||||||||||||||| |||| ||
Db    653 GCATACTCCACAGCACCTGGNTATTATTCTTGGCGAAATTCAAAGGATGGCTTCTGGNTC 712

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATG-CCGACAAGCTTGAATTTATGCACAT 926
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    713 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCCGACAAGCTTGAATTTATGCACAT 772

Qy    927 TCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTAC 986
          |||||||||||||| |  || |||||||||||| ||| |  || | ||  |
Db    773 TCTTACCCGGGTTA--CCGAANGTGGCAACAGAACTTGGTCCTCTTCCTTTGCGC---TA 827

Qy    987 TTTTCATGCAAAGAAACAG 1005
          |||||||| |||||||||
Db    828 CTTTCATGCCAAGAAACAG 846
```

```
RESULT   4
AU133051
LOCUS       AU133051                951 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  AU133051 NT2RP4 Homo sapiens cDNA clone NT2RP4001162 5', mRNA
            sequence.
ACCESSION   AU133051
VERSION     AU133051.1  GI:10993590
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 951)
  AUTHORS   Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J.,
            Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuho,Y. and
            Isogai,T.
  TITLE     HRI human cDNA project (Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y.,
            Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T., Sugano
```

```
           ,S., Masuho,Y., Isogai,T.)
  JOURNAL  Unpublished (2000)
COMMENT    Contact: Takao Isogai
           Genomics Laboratory
           Helix Research Institute
           1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
           Tel: 81-438-52-3951
           Fax: 81-438-52-3952
           Email: genomics@hri.co.jp
           HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
           Research Institute; cDNA library construction: Department of
           Virology, Institute of Medical Science, University of Tokyo, and
           Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .951
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP4001162"
                     /clone_lib="NT2RP4"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                     cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT      276 a    186 c    220 g    266 t      3 others
ORIGIN
```

Query Match 60.4%; Score 699.8; DB 9; Length 951;
Best Local Similarity 99.7%; Pred. No. 9.1e-172;
Matches 701; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       4 CGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTC 63
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17 CTAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTC 76

Qy      64 ATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCC 123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      77 ATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCC 136

Qy     124 GTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTG 183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137 GTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTG 196

Qy     184 CACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCA 243
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     197 CACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCA 256

Qy     244 GTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATG 303
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     257 GTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATG 316

Qy     304 GACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGT 363
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     317 GACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGT 376

Qy     364 ATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACA 423
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db     377 ATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTGCA 436

Qy     424 GATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAAT 483
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     437 GATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAAT 496

Qy     484 AAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGAT 543
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     497 AAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGAT 556

Qy     544 CACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATT 603
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     557 CACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATT 616

Qy     604 TTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGT 663
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     617 TTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGT 676

Qy     664 TGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGG 706
           |||||||||||||||||||||||||||||||||||||||||||
Db     677 TGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGG 719
```

```
RESULT   5
AU127381
LOCUS       AU127381                826 bp    mRNA    linear   EST 23-OCT-2000
DEFINITION  AU127381 NT2RP2 Homo sapiens cDNA clone NT2RP2001169 5', mRNA
            sequence.
ACCESSION   AU127381
VERSION     AU127381.1  GI:10952097
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 826)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .826
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP2001169"
                     /clone_lib="NT2RP2"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                     cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT      254 a    152 c    188 g    226 t      6 others
ORIGIN
```

Query Match 56.5%; Score 654.4; DB 9; Length 826;
Best Local Similarity 95.8%; Pred. No. 6.2e-160;
Matches 714; Conservative 0; Mismatches 27; Indels 4; Gaps 4;

```
Qy       4 CGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTC 63
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14 CTAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTC 73

Qy      64 ATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCC 123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74 ATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCC 133

Qy     124 GTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTG 183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     134 GTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTG 193

Qy     184 CACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCA 243
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     194 CACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCA 253

Qy     244 GTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATG 303
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     254 GTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATG 313
```

```
Qy    304 GACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGT 363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    314 GACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGT 373

Qy    364 ATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACA 423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    374 ATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACA 433

Qy    424 GATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAAT 483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    434 GATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAAT 493

Qy    484 AAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGAT 543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    494 AAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGAT 553

Qy    544 CACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCAT-GGTGAAGAAGGAATAAT 602
          |||||||||||||| || ||||||||||||||||||| |||| |||||| || |||||||
Db    554 CACAGCAAAAGGAACA-TTTTGTTTGTGTGCTTCTGAACCATGGGTGAAAAAAGAATAAT 612

Qy    603 TTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCG 662
          ||||||||||||||||||||||||||||||||||||  ||||||||||||| | ||||||
Db    613 TTTTGGAACAAATGGACCTGTTGACCTGAAAAAAAATACAAACTTTTTCA-AAGGGATCN 671

Qy    663 TTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 722
          ||||  || ||||||||||||||| ||  |||   |  | ||||||||||||||||||| |
Db    672 TTGTTAAAATCTAACTGGAAAACCAAACTTTTCATTATTCCAGGCCTGCCGTGGTACAAA 731

Qy    723 ACTGG-ACTGTGGCATTGAGACAGA 746
             || ||||||||||||| ||| |
Db    732 CTGGGAACTGTGGCATTGAAACAAA 756
```

```
RESULT   6
BE780308
LOCUS       BE780308                911 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601468289F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3871390 5',
            mRNA sequence.
ACCESSION   BE780308
VERSION     BE780308.1  GI:10201506
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 911)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9624  row: a  column: 23
            High quality sequence stop: 768.
FEATURES             Location/Qualifiers
     source          1. .911
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3871390"
                     /clone_lib="NIH_MGC_67"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
```

```
BASE COUNT      287 a    175 c    205 g    244 t
ORIGIN
```

```
  Query Match             56.2%;  Score 651.4;  DB 10;  Length 911;
  Best Local Similarity   89.7%;  Pred. No. 3.9e-159;
  Matches  769;  Conservative    0;  Mismatches   76;  Indels   12;  Gaps    6;

Qy    223 ATGGAGAACACT-GAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGAT 281
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGAGAACACTGGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGAT 60

Qy    282 CATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGA 341
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGA 120

Qy    342 TTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGG 401
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGG 180

Qy    402 AATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAA 461
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAA 240

Qy    462 CTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGAT 521
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGAT 300

Qy    522 GCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAG 581
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    301 GCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAG-TTTGTTTGTGTGCTTCTGAG 359

Qy    582 CCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAAC 641
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 CCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAAC 419

Qy    642 AAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTAT 701
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 AAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTAT 479

Qy    702 TCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGA 761
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 TCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGA 539

Qy    762 TGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTC-TTGTATGCATACTCCACAG 820
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    540 TGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTTGTATGCATACTCCACAG 599

Qy    821 CACCTGG----TTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCG 876
          |||||||    ||  |  | | | |||||  |  |  |   | ||||||||||||||||||
Db    600 CACCTGGGTTATTTATTCTATGGCCGAAAATTCACAGGATGGTCCTGGTTCATCCAGTCG 659

Qy    877 CTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACA---TTCTTACC 933
          ||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db    660 GCTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAAATTTATGCAACATTTCTTACC 719

Qy    934 CGGGTTAACCGAAAGGTGG--CAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTC 991
          |||||||||||| ||||||   | ||||||||||||||||||||||          ||| |
Db    720 CGGGTTAACCGACAGGTGGGCAACCAGAATTTGAGTCCTTTTCCTTGACGCTACTTTTCC 779

Qy    992 ATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATC 1051
          ||||| |  ||      | ||||| |  || |     || ||||    ||| ||  |
Db    780 ATGCAGAAGAAACAGATGCATGTTTGTTTCCATGCTTCCCAAAGACTCTCATTTCATCCT 839

Qy   1052 ACTAAAGAAATGGTTGG 1068
          |  | |    |||| ||
Db    840 AACACAATGTTGGTCGG 856
```

```
RESULT   7
AU127028
LOCUS       AU127028                751 bp    mRNA    linear   EST 23-OCT-2000
DEFINITION  AU127028 NT2RP2 Homo sapiens cDNA clone NT2RP2000417 5', mRNA
```

```
            sequence.
ACCESSION   AU127028
VERSION     AU127028.1  GI:10951744
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 751)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .751
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP2000417"
                     /clone_lib="NT2RP2"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                     cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT      242 a    129 c    175 g    194 t     11 others
ORIGIN
```

```
  Query Match           53.4%;  Score 618.8;  DB 9;  Length 751;
  Best Local Similarity  97.5%;  Pred. No. 1.2e-150;
  Matches  654;  Conservative    0;  Mismatches   13;  Indels    4;  Gaps    3;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     82 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 141

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    142 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 201

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    202 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 261

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 321

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    322 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 381

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    382 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 441

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db    442 TGTGTGCTTCTGANCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 501

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT-GGAAAACC 686
          |||||||||||||||||||||||||| |||||||||||| || ||||||| ||||||||
Db    502 CTGAAAAAAATAACAAACTTTTTCAGANGGGATCGTTGTANAANTCTAACTGGGAAAACC 561

Qy    687 CAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGA 746
          |||||||||||||||||| |||||||||||||||||||||||||||||||||| |||||
Db    562 CAAACTTTTCATTATTCANGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGANACAGA 621

Qy    747 CAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTA 806
          |||||||||||||||||||||||||||||||||||||||||| |||| | ||||||||||
Db    622 CAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCA-TGGAAGNCGACTTCTTGTN 680

Qy    807 TGCATACTCCACAGCACCTGGTTATTATTCTTGG--CGAAATTCAAAGGATGGCTCCTGG 864
          |||||||||||| ||||||||||||||||||||   ||||| ||||||||||||||||||
Db    681 TGCATACTCCACANCACCTGGTTATTATTCTTGGGCNGAAATCCAAAGGATGGCTCCTGG 740

Qy    865 TTCATCCAGTC 875
          ||||||||| ||
Db    741 TTCATCCANTC 751
```

```
RESULT   8
AV649560
LOCUS       AV649560                624 bp    mRNA    linear   EST 15-JAN-2002
DEFINITION  AV649560 GLC Homo sapiens cDNA clone GLCBVC04 3', mRNA sequence.
ACCESSION   AV649560
VERSION     AV649560.1  GI:9870574
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
            Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
            Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
            Hu,G., Gu,J., Chen,Z. and Han,Z.
  TITLE     Insight into hepatocellular carcinogenesis at transcriptome level
            by comparing gene expression profiles of hepatocellular carcinoma
            with those of corresponding noncancerous liver
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
  MEDLINE   21625106
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="GLCBVC04"
                     /clone_lib="GLC"
                     /tissue_type="corresponding non cancerous liver tissue"
                     /dev_stage="Adult"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      182 a    116 c    156 g    166 t      4 others
ORIGIN
```

```
  Query Match           51.9%;  Score 601;  DB 9;  Length 624;
  Best Local Similarity  98.4%;  Pred. No. 5.1e-146;
  Matches  615;  Conservative    0;  Mismatches    9;  Indels    1;  Gaps    1;

Qy      8 CGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATAC 67
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db      1 CGGATGGGTGCTATTGTGAGGCGGTTGTGGAAGAGTTTCGTGAGTGCTCGCAGCTCATAC 60

Qy     68 CTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGA 127
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     61 CTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGA 120

Qy    128 GGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACC 187
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACC 180

Qy    188 TGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGG 247
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGCCTCTTCCCCCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGG 240

Qy    248 ATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACT 307
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACT 300

Qy    308 CTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAA 367
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAA 360

Qy    368 TAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATG 427
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATG 420

Qy    428 TCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAA 487
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAA 480

Qy    488 ATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACA 547
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACA 540

Qy    548 GCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTG 607
          ||| |||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db    541 GCANAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAANTA-TTTTG 599

Qy    608 GAACAAATGGACCTGTTGACCTGAA 632
          ||||| ||||||| |||||||||
Db    600 GAACANATGGACCNTGTGACCTGAA 624


RESULT  9
AV649613
LOCUS       AV649613                624 bp    mRNA    linear   EST 15-JAN-2002
DEFINITION  AV649613 GLC Homo sapiens cDNA clone GLCBVG11 3', mRNA sequence.
ACCESSION   AV649613
VERSION     AV649613.1  GI:9870627
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
            Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
            Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
            Hu,G., Gu,J., Chen,Z. and Han,Z.
  TITLE     Insight into hepatocellular carcinogenesis at transcriptome level
            by comparing gene expression profiles of hepatocellular carcinoma
            with those of corresponding noncancerous liver
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
  MEDLINE   21625106
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="GLCBVG11"
                     /clone_lib="GLC"
                     /tissue_type="corresponding non cancerous liver tissue"
                     /dev_stage="Adult"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      181 a    116 c    157 g    163 t      7 others
ORIGIN


  Query Match           51.4%;  Score 596.2;  DB 9;  Length 624;
  Best Local Similarity   97.6%;  Pred. No. 9.1e-145;
  Matches  601;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      8 CGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATAC 67
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db      1 CGGATGGGTGCTATTGTGAGGCGGTTGTGGAAGAGTTTCGTGAGTGCTCGCAGCTCATAC 60

Qy     68 CTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGA 127
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGA 120

Qy    128 GGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACC 187
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACC 180

Qy    188 TGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGG 247
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGCCTCTTCCCCCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGG 240

Qy    248 ATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACT 307
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACT 300

Qy    308 CTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAA 367
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAA 360

Qy    368 TAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATG 427
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATG 420

Qy    428 TCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAA 487
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAA 480

Qy    488 ATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACA 547
          |||||||||||||||||||||||||||||| |||||||||||| |||| || |||||||
Db    481 ATGATCTTACACGTGAAGAAATTGTGGAATNGATGCGTGATGTTNCTAANGANGATCACA 540

Qy    548 GCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTG 607
          ||| | ||||||  ||||| |||||||||||||||||||||||||    |||||||
Db    541 GCANAGGGAGCAAGTTTGTATGTGTGCTTCTGAGCCATGGTGAAGAAGGNNATATTTTTG 600

Qy    608 GAACAAATGGACCTGT 623
          ||||||||||||||||
Db    601 GAACAAATGGACCTGT 616


RESULT  10
BG177093
LOCUS       BG177093                946 bp    mRNA    linear   EST 06-FEB-2001
DEFINITION  602313329F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4419175 5',
            mRNA sequence.
ACCESSION   BG177093
VERSION     BG177093.1  GI:12683796
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 946)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
```

```
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis Staudt, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10154  row: j  column: 08
            High quality sequence stop: 644.
FEATURES             Location/Qualifiers
     source          1. .946
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4419175"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      271 a    237 c    226 g    212 t
ORIGIN
```

```
  Query Match            50.0%;  Score 579.8;  DB 10;  Length 946;
  Best Local Similarity  92.6%;  Pred. No. 1.9e-140;
  Matches  677;  Conservative    0;  Mismatches   42;  Indels   12;  Gaps    6;

Qy      1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60
          ||  | ||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db      6 GCGTGGGCGGATGGGTGCTATTGTGAGGCGGTTGTGGAAGAGTTTCGTGAGTGCTCGCAG 65

Qy     61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     66 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 125

Qy    121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    126 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 185

Qy    181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    186 CTGCACCTGCCTCTTCCCCCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 245

Qy    241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    246 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 305

Qy    301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    306 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 365

Qy    361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    366 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 425

Qy    421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    426 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 485

Qy    481 AATAA--AAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAG 538
          ||| |   |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    486 AATCACCCAATGATCTTACACGTGAAGAAA-TGTGGAATTGATGCGTGATGTTTCTAAAG 544

Qy    539 AAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAA---GAAG 595
          ||||||||||||||||||||    | |||||||||||||||||||||    | |
Db    545 AAGATCACAGCAAAAGGAGCAGTT---GTAGTGTGCTTCTGAGCCATGGTGAACGCAACG 601

Qy    596 GAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAG 655
          ||||| | |||||||||||||||||||||||||||||||||   |||||| ||||||
Db    602 ACATAATCTCTGGAACAAATGGACCTGTTGACCTGAAAAAAATA--CAACTTTCTCAGAG 659

Qy    656 GGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTG 715
          | ||||| ||||| |||||| ||||||| || ||  |  |||| |  | | ||
Db    660 GCGATCGGTGTAG-AGTCTACCTGGAAACCCAAATTTCATCATCAGGCCGCCCGTGGTAC 718

Qy    716 GTACAGAACTG 726
          | || ||  ||
Db    719 GAACGGACGTG 729
```

```
RESULT  11
BG170404
LOCUS       BG170404                 965 bp    mRNA    linear   EST 06-FEB-2001
DEFINITION  602322734F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4425920 5',
            mRNA sequence.
ACCESSION   BG170404
VERSION     BG170404.1  GI:12677107
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 965)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10172  row: c  column: 09
            High quality sequence stop: 658.
FEATURES             Location/Qualifiers
     source          1. .965
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4425920"
                     /clone_lib="NIH_MGC_89"
                     /tissue_type="hypernephroma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      273 a    240 c    222 g    230 t
ORIGIN
```

```
  Query Match            49.7%;  Score 576.2;  DB 10;  Length 965;
  Best Local Similarity  85.9%;  Pred. No. 1.6e-139;
  Matches  721;  Conservative    0;  Mismatches  108;  Indels   10;  Gaps    7;

Qy     28 GCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTGGCTGTGTATCCGTGG 87
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCTGTGGCTGTGTATCCGTGG 60

Qy     88 CCACAGCTGGTTGGCGTCGCCTTG-AAATCCCAGGCCGTGAGGAGTTAGCGAGCCCTGCT 146
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     61 CCACAGCTGGTTGGCGTCGCCTTGCAAATCCCAGGCCGTGAGGAGTTAGCGAGCCCTGCT 120

Qy    147 CACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCTCTTCCCGCATTCTC 206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    121 CACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTGCCTCTTCCCCCATTCTC 180
```

```
Qy    207 ATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAA 266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAA 240

Qy    267 TTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAA 326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAA 300

Qy    327 CAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTT 386
          ||||||||||||||||||||||||||||||||||||||| || ||| || ||||||||||
Db    301 CAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATACTACTTACTACTAAGAATTT 360

Qy    387 TCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAG 446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAG 420

Qy    447 GGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGA 506
          |||||||||||||||||||||||||||||||| | |||||||||||||||| ||
Db    421 GGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATCACAATGATCTTACACGTGACGA 480

Qy    507 -AATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTG 565
           | | ||||||||||||||| ||| ||||    |          |||   |     | |
Db    481 CACTCGTGGAATTGATGCGTGCTGTGTCTACCAGACGATCACAGCAAACAGGAGCAGTCG 540

Qy    566 TTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTG 625
          | |||||||||||||||| ||||||  | | |  | ||| || ||||||||||||||||
Db    541 GTCGTGTGCTTCTGAGCCACGGTGAACGACGGACTAATTTCGGCACAAATGGACCTGTTG 600

Qy    626 ACC--TGAAAAAAATAACAAACTTTTTCA--GAGGGGATCGTTGTAGAAGTCTAACTGGA 681
          |||    || ||||||||||||||||||    |||  |||| ||||||||||||||||||
Db    601 ACCCTGAAACAAAATAACAAACTTTTTCACGAGGGGATTCGTAGTAGAAGTCTAACTGGA 660

Qy    682 AAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGA- 740
          |||||||  | |  |||||    |||||||||||||| ||||||  ||||||  |||| 
Db    661 AAACCCAACTTGTCAATTATCAGCGCCTGCCGTGGTACCGAACTGACTTGTGGCCTTGAC 720

Qy    741 GACAGACAGTGG-TGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACT 799
          |||  |||| || || |||||||||| |||   ||||| | ||||||  ||  || ||||
Db    721 GACGCACAGCGGCTGGTGATGATGACCTGGGTGGTCATCACATACCACCGGGCGCGGACT 780

Qy    800 TCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGC 858
            ||| ||||| | |  ||    |  | | |||| | |||   |  |||| |  || |
Db    781 --TTGAATGCATTCCCACCAAGCACATGCTTTTATACATGGGACATATCAAGGACTGCC 837
```

```
RESULT  12
BI080567
LOCUS       BI080567                878 bp    mRNA    linear   EST 20-JUN-2001
DEFINITION  602877052F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5008572 5',
            mRNA sequence.
ACCESSION   BI080567
VERSION     BI080567.1  GI:14498897
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 878)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11053  row: h  column: 13
            High quality sequence stop: 794.
```

```
FEATURES             Location/Qualifiers
     source          1. .878
                     /organism="Mus musculus"
                     /strain="FVB/N-3"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5008572"
                     /clone_lib="NCI_CGAP_Mam2"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      266 a    174 c    227 g    211 t
ORIGIN
```

```
  Query Match             47.1%;  Score 545.6;  DB 10;  Length 878;
  Best Local Similarity   80.4%;  Pred. No. 1.6e-131;
  Matches  677;  Conservative    0;  Mismatches  159;  Indels    6;  Gaps    3;

Qy    209 TAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATT 268
          ||| ||| ||  |||||||||||||   ||| ||||||||||||||||||||||| ||||
Db     27 TAAGAAAAGTGACCATGGAGAACAACAAAACCTCAGTGGATTCAAAATCCATTAATAATT 86

Qy    269 TGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 328
          | |||  ||||| |||||||||| ||| | ||| |||||||||| || |  |||||||  |
Db     87 TTGAAGTAAAGACCATACATGGGAGCAAGTCAGTGGACTCTGGGATCTATCTGGACAGTA 146

Qy    329 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTC 388
          |||| |||||||||||||||| |||||  |||| ||||||||||||||||||||| || |
Db    147 GTTACAAAATGGATTATCCTGAAATGGGCATATGCATAATAATTAATAATAAGAACTTCC 206

Qy    389 ATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGG 448
          |||| |||||||||||| |||||||| |||||||| |||| || |||| |||||||| |
Db    207 ATAAGAGCACTGGAATGTCATCTCGCTCTGGTACGGATGTGGACGCAGCCAACCTCAGAG 266

Qy    449 AAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAA 508
          | |||||||    | |||||||  |||||||||||||||||||||||||||| |||||||| |
Db    267 AGACATTCATGGGCCTGAAATACCAAGTCAGGAATAAAAATGATCTTACTCGTGAAGACA 326

Qy    509 TTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTT 568
          || ||||||| |||  |  |||||||||| |||||||| |||||||||||||| ||||| |
Db    327 TTTTGGAATTAATGGATAGTGTTTCTAAGGAAGATCATAGCAAAAGGAGCAGCTTTGTGT 386

Qy    569 GTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 628
          ||||| |||| |||||||||| |||||  | |||| ||| |||||||| |||||||| | |
Db    387 GTGTGATTCTAAGCCATGGTGATGAAGGGGTCATTTATGGGACAAATGGGCCTGTTGAAC 446

Qy    629 TGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCA 688
          ||||||| | || | ||| |||||||| ||    ||  | ||||| |||||||| || |
Db    447 TGAAAAAGTTGACTAGCTTCTTCAGAGGCGACTACTGCCGGAGTCTGACTGGAAAGCCGA 506

Qy    689 AACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACA 748
          |||| ||||| |||||||||||||| ||||| || ||||||||||||||||||||||||||
Db    507 AACTCTTCATCATTCAGGCCTGCCGGGGTACGGAGCTGGACTGTGGCATTGAGACAGACA 566

Qy    749 GTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATG 808
          ||||   |||||| || ||||| || || || || ||||| |||||||| |||||| ||||||
Db    567 GTGGGACTGATGAGGAGATGGCTTGCCAGAAGATACCGGTGGAGGCTGACTTCCTGTATG 626

Qy    809 CATACTCCACAGCACCTGGTTATTA--TTCTTGGCGAAATTCAAAGGATGGCTCCTGGTT 866
          | ||||| |||||||||||||| ||  | || |   ||||||||||| || || || ||
Db    627 CTTACTCTACAGCACCTGGTTACTAATTCCTGGAGAAAATTCAAAGGACGGGTCGTGCTT 686

Qy    867 CATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACAT 926
          |||||||| |||||   |||||||| | ||| || ||||||| ||||| ||||||||
Db    687 CATCCAGTCCCTTTGCAGCATGCTGAAGCTGTACGCGCACAAGCTAGAATTAATGCACAT 746

Qy    927 TCT-TACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTA 985
          |||  || || |||||| | |||| |||| | ||||| |||||||| ||| | ||| | |
Db    747 TCTCAACTCGCGTTAACAGGAAGGGGGCACCGGAATTCGAGTCCTTCTCCCTGGACTCCA 806
```

```
Qy    986 CTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATT 1045
          | |||| |||||||||    |   |      ||  |||||||||| || | |||| |  |
Db    807 ATGTTCAAGCAAAGAAA---AGATCCGTGTATGGGTCCATGCTCAGAACAAAACTGTCCT 863

Qy   1046 TT 1047
          ||
Db    864 TT 865


RESULT  13
BI652204
LOCUS       BI652204                794 bp    mRNA    linear   EST 12-SEP-2001
DEFINITION  603299029F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5339623 5',
            mRNA sequence.
ACCESSION   BI652204
VERSION     BI652204.1  GI:15566440
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 794)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11862  row: j  column: 08
            High quality sequence start: 37
            High quality sequence stop: 788.
FEATURES             Location/Qualifiers
     source          1. .794
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5339623"
                     /clone_lib="NIH_CGAP_Mam3"
                     /tissue_type="tumor, gross tissue"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert 2 kb. Library constructed by Life
                     Technologies, catalog #12017-018. Investigators providing
                     samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference
                     for transgenic model: Xu et al., Nature Genetics 22, 37-43
                     (1999). Note: this is a NCI_CGAP Library."
BASE COUNT      225 a    164 c    209 g    196 t
ORIGIN


  Query Match             45.1%;  Score 523.2;  DB 10;  Length 794;
  Best Local Similarity   82.1%;  Pred. No. 1.1e-125;
  Matches  614;  Conservative    0;  Mismatches  133;  Indels    1;  Gaps    1;

Qy    310 GGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATA 369
          |||   |  |||||||  ||||| | ||||||||||||||| |||||  |||| ||||||
Db     39 GGATGCTATCTGGACAGTAGTTACAGAATGGATTATCCTGAAATGGGCATATGCATAATA 98

Qy    370 ATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTC 429
          |  |||||||||| || ||||| |||||||||||| ||||||| |||||||| |||||
Db     99 AAAAATAATAAGAACTTCCATAAGAGCACTGGAATGTCATCTCGCTCTGGTACGGATGTG 158

Qy    430 GATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAAT 489
          || ||||| |||||||| || |||||||    | |||||||  |||||||||||||||||
Db    159 GACGCAGCCAACCTCAGAGAGACATTCATGGGCCTGAAATACCAAGTCAGGAATAAAAAT 218

Qy    490 GATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGC 549
          |||||||| |||||||| ||| ||||||| |||  |  |||||||| |||||||| |||
Db    219 GATCTTACTCGTGAAGACATTTTGGAATTAATGGATAGTGTTTCTAAGGAAGATCATAGC 278

Qy    550 AAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGA 609
          |||||||||| ||||| |||||| |||| |||||||||| |||||  | |||| |||
Db    279 AAAAGGAGCAGCTTTGTGTGTGTGATTCTAAGCCATGGTGATGAAGGGGTCATTTATGGG 338

Qy    610 ACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGA 669
          |||||||| |||||||| |||||||  | || | ||| |||||||| ||    ||  |
Db    339 ACAAATGGGCCTGTTGAACTGAAAAAGTTGACTAGCTTCTTCAGAGGCGACTACTGCCGG 398

Qy    670 AGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGAC 729
          ||||| |||||||| || |||| |||| |||||||||||||| ||||| || ||||||
Db    399 AGTCTGACTGGAAAGCCGAAACTCTTCATCATTCAGGCCTGCCGGGGTACGGAGCTGGAC 458

Qy    730 TGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTG 789
          ||||||||||||||||||||||||   |||||| || ||||| || || || ||||| |||
Db    459 TGTGGCATTGAGACAGACAGTGGGACTGATGAGGAGATGGCTTGCCAGAAGATACCGGTG 518

Qy    790 GAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCA 849
          ||||| |||||| |||||| |||| |||||||||||||| ||||| ||| |||||||||
Db    519 GAGGCTGACTTCCTGTATGCTTACTCTACAGCACCTGGTTACTATTCCTGGAGAAATTCA 578

Qy    850 AAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAG 909
          ||||| || || |||||||||||||| |||||   |||||||| | ||| ||  |||||
Db    579 AAGGACGGGTCGTGGTTCATCCAGTCCCTTTGCAGCATGCTGAAGCTGTACGCGCACAAG 638

Qy    910 CTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCC 969
          || |||||||||||||||| || || || |   ||| |  |||||| |||| ||||||
Db    639 CTAGAATTTATGCACATTCTCACTCGCGTAACAGGAAGGTGGGCAACGGAATTCGAGTCC 698

Qy    970 TTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGC-T 1028
          || ||| | ||| | |||| || |||||||||||||| || |||||||| |||||| |
Db    699 TTCTCCCTGGACTCCACTTTCCACGCAAAGAAACAGATCCCGTGTATTGTGTCCATGCTT 758

Qy   1029 CACAAAAGAACTCTATTTTTATCACTAA 1056
          ||| |||||||| || |||||||||| |
Db    759 CACGAAAGAACTGTACTTTTATCACTTA 786


RESULT  14
AU125557
LOCUS       AU125557                819 bp    mRNA    linear   EST 23-OCT-2000
DEFINITION  AU125557 NT2RM4 Homo sapiens cDNA clone NT2RM4001791 5', mRNA
            sequence.
ACCESSION   AU125557
VERSION     AU125557.1  GI:10950273
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 819)
  AUTHORS   Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,
            Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,
            Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki
            ,Y., Sugano,S., Isogai,T.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
```

```
     source          1. .819
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RM4001791"
                     /clone_lib="NT2RM4"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
                     precursor cells"
BASE COUNT      244 a    149 c    168 g    254 t      4 others
ORIGIN


  Query Match             45.0%;  Score 521.8;  DB 9;  Length 819;
  Best Local Similarity   99.3%;  Pred. No. 2.5e-125;
  Matches  545;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    2;

Qy     611 CAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAA 670
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAA 60

Qy     671 GTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACT 730
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACT 120

Qy     731 GTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGG 790
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGG 180

Qy     791 AGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAA 850
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAA 240

Qy     851 AGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGC 910
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGC 300

Qy     911 TTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCT 970
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCT 360

Qy     971 TTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCA 1030
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCA 420

Qy    1031 CAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGT 1090
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGT 480

Qy    1091 ATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTACT 1150
           |||||||||||||||||||||||| |||||||||||||||||||   ||||||||
Db     481 ATGCCAAGTGAGAAGATGGTATATTT-GGTACTGTATTTCCCTCTCATT-TTGACCTACT 538

Qy    1151 CTCATGCTG 1159
           |||||||||
Db     539 CTCATGCTG 547


RESULT  15
BF966141
LOCUS       BF966141                 602 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  602286412F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4375323 5',
            mRNA sequence.
ACCESSION   BF966141
VERSION     BF966141.1  GI:12333356
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 602)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10040  row: g  column: 04
            High quality sequence stop: 599.
FEATURES             Location/Qualifiers
     source          1. .602
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4375323"
                     /clone_lib="NIH_MGC_95"
                     /tissue_type="hippocampus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.5 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for  full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health).  Note: this is a NIH_MGC Library."
BASE COUNT      206 a     97 c    149 g    150 t
ORIGIN


  Query Match             42.8%;  Score 496.4;  DB 10;  Length 602;
  Best Local Similarity   98.6%;  Pred. No. 1e-118;
  Matches  511;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy     208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 145

Qy     268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 205

Qy     328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     206 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 265

Qy     388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     266 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 325

Qy     448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     326 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 385

Qy     508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     386 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTAGTT 445

Qy     568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
           |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     446 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAA-TTTTGGAACAAATGGACCTGTTGAC 504

Qy     628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
           |||||||||||||||||||||||||||||||||||||||||||||||||   ||||| ||
Db     505 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTACTGGGAAAAACC 564

Qy     688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACT 725
           ||||||||||||||||| ||||||||||||||||||
Db     565 AAACTTTTCATTATTCAGGGCTGCCGTGGTACAGAACT 602
```

Search completed: August 15, 2002, 23:02:17
Job time: 5583 sec

```
OM of: US-09-613-508-4 to: GenEmbl:*   out_format : pfs

Date: Aug 15, 2002  8:46 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=x1h
 -Q=/cgn2_1/USPTO_spool/US09613508/runat_14082002_112647_7223/app_query.fasta_1.701
 -DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
 -MINLEN=0 -MAXLEN=2000000000 -USER=US09613508_@CGN1_1_4651
 -NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-613-508-4
Query length: 277
Database: GenEmbl:*
Database sequences: 1797656
Database length: 1873333701
Search time (sec): 2733.900000


score_list:
Sequence           Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_pat:AX207737      + 1463.00 2743.07  2.0e-144  834  ! AX207737 Sequence 35 from Paten
gb_pr:HSU26943       + 1463.00 2743.07  2.0e-144  834  ! U26943 Human cysteine protease
gb_pat:AR109589      + 1463.00 2741.35  2.5e-144 1001  ! AR109589 Sequence 5 from paten
gb_pr:HSU13738       + 1463.00 2741.35  2.5e-144 1001  ! U13738 Human cysteine protease
gb_pat:AR103059      + 1463.00 2739.97  3.0e-144 1159  ! AR103059 Sequence 3 from paten
gb_pr:BC016926       + 1463.00 2732.17  8.1e-144 2652  ! BC016926 Homo sapiens, clone M
gb_pat:AR172878      + 1460.00 2726.57  1.7e-143 2635  ! AR172878 Sequence 3 from paten
gb_pat:AX134409      + 1460.00 2726.57  1.7e-143 2635  ! AX134409 Sequence 5 from Paten
gb_pat:AX331135      + 1460.00 2726.57  1.7e-143 2635  ! AX331135 Sequence 1644 from Pa
gb_pr:HSU13737       + 1460.00 2726.57  1.7e-143 2635  ! U13737 Human cysteine protease
gb_pat:AR109591      + 1453.00 2722.49  2.8e-143 1001  ! AR109591 Sequence 9 from paten
gb_pat:AX207740      + 1453.00 2704.96  2.6e-142 6436  ! AX207740 Sequence 38 from Pate
gb_pat:AX207741      + 1453.00 2704.58  2.8e-142 6703  ! AX207741 Sequence 39 from Pate
gb_pat:AX207742      + 1453.00 2704.24  2.9e-142 6946  ! AX207742 Sequence 40 from Pate
gb_pat:AX207743      + 1453.00 2703.92  3.0e-142 7189  ! AX207743 Sequence 41 from Pate
gb_pat:AX207746      + 1442.00 2683.10  4.4e-141 7248  ! AX207746 Sequence 44 from Pate
gb_pr:HSA413269      + 1441.00 2701.59  4.1e-142  834  ! AJ413269 Homo sapiens mRNA for
gb_ro:RNU49930       + 1317.00 2463.98  7.0e-129 1249  ! U49930 Rattus norvegicus ICE-l
gb_ro:RNU84410       + 1304.00 2440.83  1.4e-127 1081  ! U84410 Rattus norvegicus inter
gb_pat:AR059697      + 1304.00 2440.02  1.5e-127 1177  ! AR059697 Sequence 1 from paten
gb_pat:AR059698      + 1304.00 2440.02  1.5e-127 1177  ! AR059698 Sequence 3 from paten
gb_om:AB029345       + 1296.00 2428.19  6.9e-127  834  ! AB029345 Sus scrofa mRNA for ca
gb_ro:D86352         + 1293.00 2422.09  1.5e-126  874  ! D86352 Mouse mRNA for CPP32, cc
gb_pat:AR172885      + 1293.00 2418.37  2.4e-126 1297  ! AR172885 Sequence 10 from pate
gb_ro:MMCASP3        + 1293.00 2418.37  2.4e-126 1297  ! Y13086 M.musculus mRNA for cas
gb_ro:MMU19522       + 1293.00 2417.46  2.7e-126 1429  ! U19522 Mus musculus cysteine p
gb_ro:MMU49929       + 1293.00 2417.22  2.8e-126 1466  ! U49929 Mus musculus ICE-like c
gb_ro:CGU27463       + 1293.00 2412.59  5.1e-126 2395  ! U27463 Cricetinae gen. sp. CPP
gb_pat:AX134427      + 1271.50 2382.55  2.4e-124  786  ! AX134427 Sequence 23 from Paten
gb_ro:RNU58656       + 1216.00 2276.80  1.9e-118  884  ! U58656 Rattus norvegicus cystei
gb_pat:AX235203      + 1143.50 2126.83  4.2e-110 3614  ! AX235203 Sequence 1 from Paten
gb_htg:AC094580      + 1124.50 2056.25  3.6e-106 144682  ! AC094580 Rattus norvegicus c
gb_htg:AC096596      + 1124.50 2054.69  4.4e-106 170674  ! AC096596 Rattus norvegicus c
gb_ro:MMU63720       + 1040.00 1939.30  1.2e-99  1609  ! U63720 Mus musculus CPP32 apop
gb_om:AF068837       + 1001.00 1874.09  5.0e-96   665  ! AF068837 Ovis aries caspase-3 m
gb_ro:RNU34685       +  996.00 1865.04  1.6e-95   639  ! U34685 Rattus norvegicus cystei
gb_ov:AF083029       +  958.00 1785.58  4.3e-91  1464  ! AF083029 Gallus gallus caspase
gb_sy:AF052645       +  912.00 1703.72  1.6e-86   873  ! AF052645 Synthetic construct re
gb_pat:AX260262      +  900.00 1674.96  6.2e-85  1673  ! AX260262 Sequence 27 from Pate
gb_pat:AX260259      +  900.00 1670.93  1.0e-84  2566  ! AX260259 Sequence 24 from Pate
gb_sy:AF052647       +  896.00 1673.71  7.3e-85   858  ! AF052647 Synthetic construct
gb_pr:AL592104       -  823.50 1489.61  1.3e-74 131489  ! AL592104 Human DNA sequen
gb_htg:AL355524      -  823.50 1485.78  2.1e-74 197371  ! AL355524 Homo sapiens chr
gb_ov:D89784         +  786.50 1466.99  2.4e-73   882  ! D89784 Xenopus laevis mRNA f
gb_ov:AB047003       +  782.00 1458.86  6.8e-73   849  ! AB047003 Danio rerio mRNA fo

seq_name: gb_pat:AX207737

seq_documentation_block:
LOCUS       AX207737               834 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 35 from Patent WO0157242.
ACCESSION   AX207737
VERSION     AX207737.1  GI:15422421
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 834)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 35 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .834
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      274 a    147 c    179 g    234 t
ORIGIN


alignment_scores:
            Quality: 1463.00              Length:    277
              Ratio:  5.282                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AX207737   ..

 Align seg 1/1  to: AX207737  from: 1  to: 834

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 50

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 100

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 150

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 200

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 250

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 300

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTGTGT 350

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 400
```

```
     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 450

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 500

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 550

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     551 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 600

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     601 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 650

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     651 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 700

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     701 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 750

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     751 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 800

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     801 CATGCTCACAAAAGAACTCTATTTTTATCAC 831

seq_name: gb_pr:HSU26943

seq_documentation_block:
LOCUS       HSU26943                834 bp    mRNA    linear   PRI 29-AUG-1995
DEFINITION  Human cysteine protease Yama mRNA, complete cds.
ACCESSION   U26943
VERSION     U26943.1  GI:857568
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 834)
  AUTHORS   Tewari,M., Quan,L.T., O'Rourke,K., Desnoyers,S., Zeng,Z.,
            Beidler,D.R., Poirier,G.G., Salvesen,G.S. and Dixit,V.M.
  TITLE     Yama/CPP32 beta, a mammalian homolog of CED-3, is a
            CrmA-inhibitable protease that cleaves the death substrate
            poly(ADP-ribose) polymerase
  JOURNAL   Cell 81 (5), 801-809 (1995)
  MEDLINE   95292347
REFERENCE   2  (bases 1 to 834)
  AUTHORS   Tewari,M., O'Rourke,K., Lin,Z. and Dixit,V.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-1995) Muneesh Tewari, Pathology, University of
            Michigan, 1301 Catherine, Ann Arbor, MI 48109, USA
FEATURES             Location/Qualifiers
     source          1. .834
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="umbilical vein endothelial cells treated with
                     tumor necrosis factor and cycloheximide"
     CDS             1. .834
                     /note="cysteine protease CPP32beta; CED-3/ICE protein
                     family"
                     /codon_start=1
                     /product="Yama protein"
                     /protein_id="AAA74929.1"
                     /db_xref="GI:857569"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVEADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
BASE COUNT      274 a    147 c    179 g    234 t
ORIGIN


alignment_scores:
            Quality: 1463.00              Length:    277
              Ratio:  5.282                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x HSU26943   ..

 Align seg 1/1  to: HSU26943  from: 1  to: 834

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 50

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 100

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 150

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 200

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 250

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 300

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 350

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 400

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 450

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 500

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 550

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     551 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 600

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     601 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 650
```

```
     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     651 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 700

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     701 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 750

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     751 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 800

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     801 CATGCTCACAAAAGAACTCTATTTTTATCAC 831

seq_name: gb_pat:AR109589

seq_documentation_block:
LOCUS       AR109589                1001 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 5 from patent US 6114132.
ACCESSION   AR109589
VERSION     AR109589.1  GI:12825865
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1001)
  AUTHORS   Desmarais,S., Gresser,M., Friesen,R., Kennedy,B., Nicholson,D.,
            Skorey,K., Ramachandran,C. and Ford-Hutchinson,A.
  TITLE     Prosphatase binding assay
  JOURNAL   Patent: US 6114132-A 5 05-SEP-2000;
FEATURES             Location/Qualifiers
     source          1. .1001
                     /organism="unknown"
BASE COUNT      310 a    171 c    227 g    293 t
ORIGIN


alignment_scores:
            Quality: 1463.00             Length:    277
              Ratio:  5.282                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AR109589   ..

 Align seg 1/1  to: AR109589  from: 1  to: 1001

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 563

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: gb_pr:HSU13738

seq_documentation_block:
LOCUS       HSU13738                1001 bp    mRNA    linear   PRI 14-APR-1995
DEFINITION  Human cysteine protease CPP32 isoform beta mRNA, complete cds.
ACCESSION   U13738
VERSION     U13738.1  GI:561667
KEYWORDS    cysteine protease; interleukin 1-beta converting enzyme; apoptotic
            protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1001)
  AUTHORS   Fernandes-Alnemri,T., Litwack,G. and Alnemri,E.S.
  TITLE     CPP32, a novel human apoptotic protein with homology to
            Caenorhabditis elegans cell death protein Ced-3 and mammalian
            interleukin-1 beta-converting enzyme
  JOURNAL   J. Biol. Chem. 269 (49), 30761-30764 (1994)
  MEDLINE   95074098
REFERENCE   2  (bases 1 to 1001)
  AUTHORS   Fernandes-Alnemri,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-AUG-1994) Teresa Fernandes-Alnemri, Pharmacology,
            Thomas Jefferson University, Jefferson Cancer Institute, 233, S.
            Tenth Street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .1001
                     /organism="Homo sapiens"
```

```
                     /db_xref="taxon:9606"
                     /clone="J6.1 Pst I fragment"
                     /cell_line="Jurkat"
                     /cell_type="T-lymphocyte"
                     /clone_lib="Human Jurkat T-lymphocyte lambda Uni-ZAP XR
                     cDNA library"
     5'UTR           1. .63
                     /note="An alternatively spliced isoform which lacks most
                     of the 5' nontranslated sequence"
     CDS             64. .897
                     /note="new member of the interleukin 1-beta converting
                     enzyme gene family of cysteine proteases"
                     /codon_start=1
                     /product="cysteine protease CPP32 isoform beta"
                     /protein_id="AAB60355.1"
                     /db_xref="GI:561668"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVEADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
     misc_difference 234
                     /note="differs from isoform alpha by one bp substitution
                     at position 234"
                     /replace="a"
     misc_difference 633
                     /note="differs from isoform alpha by one bp substitution
                     at position 633"
                     /replace="t"
BASE COUNT      310 a    171 c    227 g    293 t
ORIGIN


alignment_scores:
            Quality: 1463.00             Length:    277
              Ratio:  5.282               Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x HSU13738   ..

 Align seg 1/1  to: HSU13738  from: 1  to: 1001

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     464 TTGACCTGAAAAAAATAACAAACTTTTCAGAGGGGATCGTTGTAGAAGT 513

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 563

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: gb_pat:AR103059

seq_documentation_block:
LOCUS       AR103059               1159 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 3 from patent US 6087150.
ACCESSION   AR103059
VERSION     AR103059.1  GI:12814647
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1159)
  AUTHORS   He,W.W., Rosen,C.A., Hudson,P.L. and Hastings,G.A.
  TITLE     Interleukin-1 .beta. converting enzyme like apoptosis protease-3
            and 4
  JOURNAL   Patent: US 6087150-A 3 11-JUL-2000;
FEATURES             Location/Qualifiers
     source          1. .1159
                     /organism="unknown"
BASE COUNT      329 a    220 c    276 g    334 t
ORIGIN


alignment_scores:
            Quality: 1463.00             Length:    277
              Ratio:  5.282               Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AR103059   ..

 Align seg 1/1  to: AR103059  from: 1  to: 1159

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
```

```
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   223 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 272

    17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   273 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 322

    34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   323 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 372

    51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   373 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 422

    67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   423 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 472

    84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   473 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 522

   101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   523 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 572

   117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   573 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 622

   134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   623 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 672

   151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   673 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 722

   167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   723 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 772

   184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   773 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 822

   201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   823 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 872

   217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   873 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 922

   234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   923 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 972

   251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   973 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1022

   267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
     |||||||||||||||||||||||||||||||
  1023 CATGCTCACAAAAGAACTCTATTTTTATCAC 1053

seq_name: gb_pr:BC016926

seq_documentation_block:
LOCUS       BC016926                2652 bp    mRNA    linear   PRI 09-NOV-2001
DEFINITION  Homo sapiens, clone MGC:21241 IMAGE:4419175, mRNA, complete cds.
ACCESSION   BC016926
VERSION     BC016926.1  GI:16877343
KEYWORDS    MGC.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2652)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 27 Row: j Column: 23
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 14790118.
FEATURES             Location/Qualifiers
     source          1. .2652
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="MGC:21241 IMAGE:4419175"
                     /tissue_type="Lymph, lymphoma"
                     /clone_lib="NIH_MGC_85"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             228. .1061
                     /codon_start=1
                     /product="Unknown (protein for MGC:21241)"
                     /protein_id="AAH16926.1"
                     /db_xref="GI:16877344"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVEADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
BASE COUNT      824 a    466 c    576 g    786 t
ORIGIN


alignment_scores:
            Quality: 1463.00              Length:    277
              Ratio:  5.282                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x BC016926   ..

 Align seg 1/1  to: BC016926  from: 1  to: 2652

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     228 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 277

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     278 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 327
```

```
    34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   328 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 377

    51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   378 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 427

    67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   428 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 477

    84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   478 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 527

   101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   528 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 577

   117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   578 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 627

   134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   628 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 677

   151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   678 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 727

   167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   728 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 777

   184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   778 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 827

   201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   828 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 877

   217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   878 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 927

   234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   928 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 977

   251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   978 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1027

   267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
       |||||||||||||||||||||||||||||||
  1028 CATGCTCACAAAAGAACTCTATTTTTATCAC 1058

seq_name: gb_pat:AR172878

seq_documentation_block:
LOCUS       AR172878               2635 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 3 from patent US 6303374.
ACCESSION   AR172878
VERSION     AR172878.1  GI:17912369
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2635)
  AUTHORS   Zhang,H. and Cowsert,L.M.
  TITLE     Antisense modulation of caspase 3 expression
  JOURNAL   Patent: US 6303374-A 3 16-OCT-2001;
FEATURES             Location/Qualifiers
     source          1. .2635
                     /organism="unknown"
BASE COUNT      815 a    465 c    569 g    786 t
ORIGIN


alignment_scores:
           Quality: 1460.00              Length:    277
             Ratio:  5.271                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.639

alignment_block:
 US-09-613-508-4 x AR172878   ..

 Align seg 1/1  to: AR172878  from: 1  to: 2635

     1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   225 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 274

    17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   275 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 324

    34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   325 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

    51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 424

    67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   425 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 474

    84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   475 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 524

   101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   525 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 574

   117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   575 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 624

   134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   625 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 674

   151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 724

   167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   725 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 774

   184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
       ||||||||||||||||||:::||||||||||||||||||||||||||||
   775 GTCATAAAATACCAGTGGATGCCGACTTCTTGTATGCATACTCCACAGCA 824

   201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   825 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 874

   217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
       ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     875 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 924

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     925 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 974

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     975 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1024

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
    1025 CATGCTCACAAAAGAACTCTATTTTTATCAC 1055

seq_name: gb_pat:AX134409

seq_documentation_block:
LOCUS       AX134409               2635 bp    DNA     linear   PAT 29-MAY-2001
DEFINITION  Sequence 5 from Patent WO0129232.
ACCESSION   AX134409
VERSION     AX134409.1  GI:14271055
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2635)
  AUTHORS   Cordell,B. and Li,Y.
  TITLE     Functional cloning of genes encoding proteins/enzymes involved in
            proteolytic cleavage
  JOURNAL   Patent: WO 0129232-A 5 26-APR-2001;
            Scios Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2635
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             225. .1058
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC39521.1"
                     /db_xref="GI:14271056"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVDADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
BASE COUNT      815 a    465 c    569 g    786 t
ORIGIN


alignment_scores:
            Quality: 1460.00            Length:    277
              Ratio:  5.271               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.639

alignment_block:
 US-09-613-508-4 x AX134409   ..

 Align seg 1/1  to: AX134409  from: 1  to: 2635

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     225 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 274

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     275 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 324

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     325 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 424

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     425 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 474

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     475 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 524

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     525 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 574

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     575 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 624

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     625 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 674

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 724

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     725 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 774

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||:::||||||||||||||||||||||||||||||
     775 GTCATAAAATACCAGTGGATGCCGACTTCTTGTATGCATACTCCACAGCA 824

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     825 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 874

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     875 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 924

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     925 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 974

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     975 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1024

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
    1025 CATGCTCACAAAAGAACTCTATTTTTATCAC 1055

seq_name: gb_pat:AX331135

seq_documentation_block:
LOCUS       AX331135               2635 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 1644 from Patent WO0194629.
ACCESSION   AX331135
VERSION     AX331135.1  GI:18121769
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
```

```
  JOURNAL   Patent: WO 0194629-A 1644 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .2635
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      815 a    465 c    569 g    786 t
ORIGIN


alignment_scores:
            Quality: 1460.00              Length:    277
              Ratio:  5.271                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.639

alignment_block:
 US-09-613-508-4 x AX331135   ..

 Align seg 1/1  to: AX331135  from: 1  to: 2635

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     225 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 274

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     275 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 324

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     325 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 424

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     425 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 474

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     475 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 524

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     525 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 574

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     575 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 624

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     625 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 674

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 724

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     725 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 774

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||:::||||||||||||||||||||||||||||||||
     775 GTCATAAAATACCAGTGGATGCCGACTTCTTGTATGCATACTCCACAGCA 824

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     825 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 874

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     875 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 924

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     925 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 974

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     975 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1024

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||
    1025 CATGCTCACAAAAGAACTCTATTTTTATCAC 1055

seq_name: gb_pr:HSU13737

seq_documentation_block:
LOCUS       HSU13737                2635 bp    mRNA    linear   PRI 14-APR-1995
DEFINITION  Human cysteine protease CPP32 isoform alpha mRNA, complete cds.
ACCESSION   U13737
VERSION     U13737.1  GI:561665
KEYWORDS    cysteine protease; interleukin 1-beta converting enzyme; apoptotic
            protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2635)
  AUTHORS   Fernandes-Alnemri,T., Litwack,G. and Alnemri,E.S.
  TITLE     CPP32, a novel human apoptotic protein with homology to
            Caenorhabditis elegans cell death protein Ced-3 and mammalian
            interleukin-1 beta-converting enzyme
  JOURNAL   J. Biol. Chem. 269 (49), 30761-30764 (1994)
  MEDLINE   95074098
REFERENCE   2  (bases 1 to 2635)
  AUTHORS   Fernandes-Alnemri,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-AUG-1994) Teresa Fernandes-Alnemri, Pharmacology,
            Thomas Jefferson University, Jefferson Cancer Institute, 233, S.
            Tenth Street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .2635
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="J1.1 EcoRI/XhoI fragment"
                     /cell_line="Jurkat"
                     /cell_type="T-lymphocyte"
                     /clone_lib="Human Jurkat T-lymphocyte lambda Uni-ZAP XR
                     cDNA library"
     CDS             225. .1058
                     /note="new member of the interleukin 1-beta converting
                     enzyme gene family of cysteine proteases"
                     /codon_start=1
                     /product="cysteine protease CPP32 isoform alpha"
                     /protein_id="AAA65015.1"
                     /db_xref="GI:561666"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVDADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
     polyA_site      2635
                     /note="19 A residues"
BASE COUNT      815 a    465 c    569 g    786 t
ORIGIN


alignment_scores:
            Quality: 1460.00              Length:    277
              Ratio:  5.271                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.639
```

```
alignment_block:
 US-09-613-508-4 x HSU13737   ..

 Align seg 1/1  to: HSU13737  from: 1  to: 2635

      1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    225 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 274

     17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    275 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 324

     34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    325 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

     51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 424

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    425 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 474

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    475 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 524

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    525 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 574

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    575 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 624

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    625 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 674

    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 724

    167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    725 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 774

    184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
        ||||||||||||||||:::||||||||||||||||||||||||||||||||
    775 GTCATAAAATACCAGTGGATGCCGACTTCTTGTATGCATACTCCACAGCA 824

    201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    825 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 874

    217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    875 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 924

    234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    925 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 974

    251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    975 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1024

    267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
        ||||||||||||||||||||||||||||||
   1025 CATGCTCACAAAAGAACTCTATTTTTATCAC 1055
```

```
seq_name: gb_pat:AR109591

seq_documentation_block:
LOCUS       AR109591                1001 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 9 from patent US 6114132.
ACCESSION   AR109591
VERSION     AR109591.1  GI:12825867
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1001)
  AUTHORS   Desmarais,S., Gresser,M., Friesen,R., Kennedy,B., Nicholson,D.,
            Skorey,K., Ramachandran,C. and Ford-Hutchinson,A.
  TITLE     Prosphatase binding assay
  JOURNAL   Patent: US 6114132-A 9 05-SEP-2000;
FEATURES             Location/Qualifiers
     source          1. .1001
                     /organism="unknown"
BASE COUNT      310 a    172 c    226 g    293 t
ORIGIN


alignment_scores:
           Quality: 1453.00              Length:    277
             Ratio:  5.264                Gaps:      0
 Percent Similarity:  99.639   Percent Identity: 99.639

alignment_block:
 US-09-613-508-4 x AR109591   ..

 Align seg 1/1  to: AR109591  from: 1  to: 1001

      1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

     17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

     34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

     51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        ||||||||||||||||||||||||||||||||||  ||||||||||||
    514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTCCCGTGGTACAGA 563
```

```
     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: gb_pat:AX207740

seq_documentation_block:
LOCUS       AX207740               6436 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 38 from Patent WO0157242.
ACCESSION   AX207740
VERSION     AX207740.1  GI:15422424
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 6436)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 38 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6436
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-1XUb-C3 construct"
BASE COUNT     1577 a   1606 c   1607 g   1646 t
ORIGIN


alignment_scores:
            Quality: 1453.00             Length:    275
              Ratio:  5.284                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AX207740   ..

 Align seg 1/1  to: AX207740  from: 1  to: 6436

       3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1144 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAA 1193

      19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1194 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1243

      36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1244 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 1293

      53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1294 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 1343

      69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1344 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 1393

      86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1394 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1443

     103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1444 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 1493

     119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1494 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 1543

     136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1544 TGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 1593

     153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1594 GGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 1643

     169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1644 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 1693

     186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1694 AAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 1743

     203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1744 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 1793

     219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1794 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 1843

     236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1844 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 1893

     253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1894 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 1943

     269 uThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||
    1944 CACAAAAGAACTCTATTTTTATCAC 1968

seq_name: gb_pat:AX207741

seq_documentation_block:
LOCUS       AX207741               6703 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 39 from Patent WO0157242.
ACCESSION   AX207741
VERSION     AX207741.1  GI:15422425
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 6703)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
```

```
  JOURNAL   Patent: WO 0157242-A 39 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6703
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-2XUb-C3 construct"
BASE COUNT     1653 a   1677 c   1674 g   1699 t
ORIGIN


alignment_scores:
            Quality: 1453.00              Length:    275
              Ratio:  5.284                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AX207741   ..

 Align seg 1/1  to: AX207741  from: 1  to: 6703

       3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1411 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAA 1460

      19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1461 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1510

      36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1511 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 1560

      53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1561 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 1610

      69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1611 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 1660

      86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1661 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1710

     103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1711 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 1760

     119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1761 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 1810

     136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1811 TGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 1860

     153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1861 GGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 1910

     169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1911 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 1960

     186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1961 AAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2010

     203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    2011 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 2060

     219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    2061 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 2110

     236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    2111 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 2160

     253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    2161 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 2210

     269 uThrLysGluLeuTyrPheTyrHis 277
         |||||||||||||||||||||||||
    2211 CACAAAAGAACTCTATTTTTATCAC 2235

seq_name: gb_pat:AX207742

seq_documentation_block:
LOCUS       AX207742               6946 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 40 from Patent WO0157242.
ACCESSION   AX207742
VERSION     AX207742.1  GI:15422426
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 6946)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 40 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6946
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-3XUb-C3 construct"
BASE COUNT     1723 a   1742 c   1735 g   1746 t
ORIGIN


alignment_scores:
            Quality: 1453.00              Length:    275
              Ratio:  5.284                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AX207742   ..

 Align seg 1/1  to: AX207742  from: 1  to: 6946

       3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1654 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAA 1703

      19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1704 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1753

      36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1754 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 1803

      53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1804 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 1853

      69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1854 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 1903

      86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
         |||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
    1904 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1953

     103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1954 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 2003

     119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2004 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 2053

     136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2054 TGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 2103

     153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2104 GGAAAACCCAAACTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 2153

     169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2154 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 2203

     186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2204 AAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2253

     203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2254 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 2303

     219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2304 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 2353

     236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2354 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 2403

     253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2404 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 2453

     269 uThrLysGluLeuTyrPheTyrHis 277
         |||||||||||||||||||||||||
    2454 CACAAAAGAACTCTATTTTTATCAC 2478

seq_name: gb_pat:AX207743

seq_documentation_block:
LOCUS       AX207743               7189 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 41 from Patent WO0157242.
ACCESSION   AX207743
VERSION     AX207743.1  GI:15422427
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 7189)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 41 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7189
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-4XUb-C3 construct"
BASE COUNT     1793 a   1807 c   1796 g   1793 t
ORIGIN


alignment_scores:
            Quality: 1453.00            Length:    275
```

```
            Ratio:  5.284                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AX207743   ..

 Align seg 1/1  to: AX207743  from: 1  to: 7189

       3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1897 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAA 1946

      19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1947 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1996

      36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1997 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 2046

      53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2047 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 2096

      69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2097 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 2146

      86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2147 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 2196

     103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2197 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 2246

     119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2247 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 2296

     136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2297 TGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 2346

     153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2347 GGAAAACCCAAACTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 2396

     169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2397 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 2446

     186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2447 AAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2496

     203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2497 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 2546

     219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2547 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 2596

     236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2597 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 2646

     253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2647 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 2696

     269 uThrLysGluLeuTyrPheTyrHis 277
```

```
                    ||||||||||||||||||||||||
2697 CACAAAAGAACTCTATTTTTATCAC 2721
```

```
OM of: US-09-613-508-4 to: N_Geneseq_032802:*   out_format : pfs

Date: Aug 15, 2002  8:52 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO_spool/US09613508/runat_14082002_112648_7265/app_query.fasta_1.701
 -DB=N_Geneseq_032802 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US09613508_@CGN1_1_190 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-613-508-4
Query length: 277
Database: N_Geneseq_032802:*
Database sequences: 1736436
Database length: 858457221
Search time (sec): 267.890000


score_list:
Sequence                Strd Orig    ZScore    EScore Len  ! Documentation  ..
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13075 + 1463.00 3191.36  2.1e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT33567 + 1463.00 3187.60  3.3e-16
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV05471 + 1463.00 3187.60  3.3e-16
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT15277 + 1463.00 3187.59  3.3e-16
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAC64043 + 1463.00 3178.60  1.1e-16
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81219 + 1460.00 3184.75  4.8e-16
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAS10503 + 1460.00 3171.58  2.6e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03910 + 1460.00 3171.58  2.6e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV32608 + 1453.00 3167.27  4.5e-16
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV32615 + 1453.00 3167.27  4.5e-16
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13078 + 1453.00 3145.94  7.0e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13079 + 1453.00 3145.48  7.4e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13080 + 1453.00 3145.07  7.8e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13081 + 1453.00 3144.67  8.2e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13084 + 1442.00 3120.39  1.9e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV13959 + 1304.00 2837.66  1.0e-14
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV13960 + 1304.00 2837.66  1.0e-14
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAI66510 + 1296.00 2824.01  6.0e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAS10510 + 1293.00 2812.35  2.7e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03918 + 1271.50 2770.80  5.5e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAH26878 + 1143.50 2471.75  2.5e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81214 +  912.00 1978.80  7.1e-10
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81215 +  896.00 1943.80  6.4e-10
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81223 +  725.00 1566.94  6.2e-79
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT66992 +  725.00 1556.31  2.4e-78
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03914 +  725.00 1556.31  2.4e-7
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT66970 +  725.00 1556.15  2.5e-78
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT15276 +  712.00 1533.69  4.4e-77
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT66993 +  587.50 1253.52  1.8e-61
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81222 +  525.50 1128.49  1.6e-54
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT43537 +  517.50 1104.48  3.6e-53
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03913 +  517.50 1104.48  3.6e-5
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAZ98834 +  516.50 1105.34  3.2e-53
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAZ98839 +  516.50 1105.13  3.3e-53
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL03125 +  502.50 1073.77  1.8e-5
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL03124 +  502.50 1062.90  7.4e-5
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAV80061 +  470.00  993.77  5.3e-47
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL09303 +  469.00  990.15  8.4e-4
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL09302 +  448.00  937.06  7.6e-4
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL06952 +  448.00  933.11  1.3e-4
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD12324 +  435.00  923.42  4.4
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAZ56946 +  423.00  903.19  5.8e
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL26207 +  422.00  900.26  8.5
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV53210 +  421.00  894.31  1.8e
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81224 +  421.00  892.60  2.3e

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13075

seq_documentation_block:
ID   AAD13075 standard; cDNA; 834 BP.
XX
AC   AAD13075;
XX
DT   16-OCT-2001  (first entry)
XX
DE   Mammalia caspase-3 reporter moiety cDNA.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance;
KW   caspase-3; ss.
XX
OS   Mammalia.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain  -
XX
PS   Example 4; Page 131; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is mammalia caspase-3 cDNA which is used as a reporter moiety in
CC   the exemplification of the invention.
XX
SQ   Sequence 834 BP; 274 A; 147 C; 179 G; 234 T; 0 other;


alignment_scores:
            Quality: 1463.00              Length:    277
              Ratio:  5.282                Gaps:      0
 Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AAD13075   ..
```

```
Align seg 1/1  to: AAD13075  from: 1  to: 834

      1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      1 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 50

     17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 100

     34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    101 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 150

     51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    151 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 200

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    201 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 250

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    251 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 300

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    301 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 350

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    351 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 400

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    401 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 450

    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    451 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 500

    167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    501 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 550

    184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    551 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 600

    201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    601 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 650

    217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    651 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 700

    234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    701 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 750

    251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    751 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 800

    267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
        ||||||||||||||||||||||||||||||
    801 CATGCTCACAAAAGAACTCTATTTTTATCAC 831

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT33567

seq_documentation_block:

ID   AAT33567 standard; cDNA; 1158 BP.
XX
AC   AAT33567;
XX
DT   06-DEC-1996  (first entry)
XX
DE   Pro-Yama cDNA.
XX
KW   Yama; CrmA; apoptosis; Fas receptor; T-lymphocyte; AIDS; HIV;
KW   human immunodeficiency virus; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             224..1057
FT                   /*tag=  a
XX
PN   WO9625501-A2.
XX
PD   22-AUG-1996.
XX
PF   09-FEB-1996;   96WO-US01882.
XX
PR   01-JUN-1995;   95US-0457731.
PR   13-FEB-1995;   95US-0389812.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Dixit VM;
XX
DR   WPI; 1996-393402/39.
DR   P-PSDB; AAW00677.
XX
PT   New nucleic acid encoding Yama protein or CrmA mutant - useful for
PT   modulating apoptosis, maintaining T cell viability in AIDS patients
PT   and for drug screening
XX
PS   Disclosure; Fig 1; 118pp; English.
XX
CC   A cDNA clone (AAT33567) codes for pro-Yama (AAW00677), a zymogen which
CC   upon activation cleaves PARP to an 85 kDa form.  Activated Yama has
CC   the ability to modulate cellular function associated with the Fas
CC   receptor pathway such as Fas-associated apoptosis.  This activity is
CC   inhibited by CrmA (see also AAW00678).  The cDNA was isolated from a
CC   cDNA library prepd. from human umbilical vein endothelial cells
CC   treated with tumour necrosis factor and cycloheximide.  It can be
CC   used to produce Yama in a host cell for use in drug screening.
CC   Yama and anti-Yama protein nucleic acids, or antisense RNA, can be
CC   used to modulate cellular functions such as apoptosis and immune
CC   disorders mediated by the Fas pathway.
XX
SQ   Sequence 1158 BP; 328 A; 220 C; 275 G; 335 T; 0 other;

alignment_scores:
            Quality: 1463.00              Length:    277
              Ratio:  5.282                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AAT33567    ..

 Align seg 1/1  to: AAT33567  from: 1  to: 1158

      1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    224 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 273

     17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    274 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 323

     34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 324 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 373

  51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 374 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 423

  67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 424 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 473

  84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 474 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 523

 101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 524 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 573

 117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 574 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 623

 134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 624 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 673

 151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 674 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 723

 167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 724 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 773

 184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 774 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 823

 201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 824 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 873

 217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 874 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 923

 234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 924 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 973

 251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 974 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1023

 267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
     ||||||||||||||||||||||||||||||
1024 CATGCTCACAAAAGAACTCTATTTTTATCAC 1054
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV05471

seq_documentation_block:

```
ID   AAV05471 standard; DNA; 1158 BP.
XX
AC   AAV05471;
XX
DT   02-JUL-1998  (first entry)
XX
DE   Nucleic acid encoding a protein designated YAMA.
XX
KW   CrmA; tumour necrosis factor receptor; TNF-R; U1-70; inhibition;
KW   apoptosis; activation; U1-70 path; apoptosis path; identification;
KW   modulation; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             224..1057
FT                   /*tag= a
XX
PN   JP09299077-A.
XX
PD   25-NOV-1997.
XX
PF   27-JAN-1997;   97JP-0012932.
XX
PR   09-FEB-1996;   96US-0591605.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
DR   WPI; 1998-056551/06.
DR   P-PSDB; AAW41688.
XX
PT   Controlling apoptosis by inhibiting activation of the U1-70 pathway
PT   - by introduction of nucleic acid encoding protein with CrmA
PT   biological activity into the cell
XX
PS   Disclosure; Fig 1; 48pp; Japanese.
XX
CC   The present sequence encodes a non-naturally occurring protein designated
CC   YAMA. The specification describes a novel method of controlling cell
CC   functions controlled by the tumour necrosis factor receptor (TNF-R)
CC   pathway containing U1-70. The method comprises introducing a nucleic
CC   acid molecule encoding CrmA into a cell, and culturing the cell under
CC   suitable conditions for transcription and translation of the CrmA nucleic
CC   acid. Method for preventing and inhibiting apoptosis by inhibiting
CC   activation of the U1-70 path, identifying the chemical related to the
CC   apoptosis path, and screening a candidate chemical having biological
CC   function in the apoptosis path are also described. The methods can be
CC   used to identify substances which modulate apoptosis in cells controlled
CC   by the TNF-R pathway.
XX
SQ   Sequence 1158 BP; 328 A; 220 C; 275 G; 335 T; 0 other;
```

alignment_scores:

```
           Quality: 1463.00              Length:    277
             Ratio:  5.282                 Gaps:      0
Percent Similarity: 100.000   Percent Identity: 100.000
```

alignment_block:

US-09-613-508-4 x AAV05471 ..

Align seg 1/1 to: AAV05471 from: 1 to: 1158

```
   1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 224 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 273

  17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 274 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 323

  34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 324 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 373

  51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 374 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 423

  67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 424 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 473

  84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
```

```
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 474 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 523

 101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 524 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 573

 117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 574 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 623

 134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 624 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 673

 151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 674 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 723

 167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 724 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 773

 184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 774 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 823

 201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 824 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 873

 217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 874 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 923

 234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 924 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 973

 251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 974 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1023

 267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
     ||||||||||||||||||||||||||||||
1024 CATGCTCACAAAAGAACTCTATTTTTATCAC 1054
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT15277

seq_documentation_block:

```
ID   AAT15277 standard; cDNA; 1159 BP.
XX
AC   AAT15277;
XX
DT   28-OCT-1996  (first entry)
XX
DE   Human interleukin-I-converting enzyme-like apoptosis protease-4 cDNA.
XX
KW   ICE-LAP-4; interleukin-I-converting enzyme-like apoptosis protease;
KW   enzyme; Alzheimer's disease; Parkinson's disease; septic shock;
KW   head injury; rheumatoid arthritis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             115..1137
FT                   /*tag=  a
XX
PN   WO9613603-A1.
XX
PD   09-MAY-1996.
XX
PF   06-JUN-1995;   95WO-US07235.
XX
PR   01-NOV-1994;   94US-0334251.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Hastings GA,  He W,  Hudson PL,  Rosen CA;
XX
DR   WPI; 1996-239509/24.
DR   P-PSDB; AAR95831.
XX
PT   Human ICE-LAP-3 and -4 DNA and protein  - useful in the diagnosis
PT   and treatment of Alzheimer's disease, Parkinson's disease,
PT   rheumatoid arthritis, septic shock and head injury
XX
PS   Claim 1; Page 44; 67pp; English.
XX
CC   This sequence was derived from a cDNA library from human endometrial
CC   tumuor and may be used for the recombinant production of the ICE-LAP-
CC   4 protein e.g. in COS cells.  The encoded protein may be used
CC   therapeutically, e.g. as an antitumor or antiviral agent and to
CC   control embryonic development and tissue homeostasis.  The protein
CC   can also be used to treat immunosuppression disorders, such as AIDS,
CC   by targeting virus infected cells for cell death.  The DNA may find
CC   use in gene therapy applications.
XX
SQ   Sequence 1159 BP; 329 A; 220 C; 276 G; 334 T; 0 other;
```

alignment_scores:

```
            Quality: 1463.00              Length:    277
              Ratio:  5.282                Gaps:      0
Percent Similarity: 100.000    Percent Identity: 100.000
```

alignment_block:

US-09-613-508-4 x AAT15277 ..

Align seg 1/1 to: AAT15277 from: 1 to: 1159

```
   1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 223 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 272

  17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 273 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 322

  34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 323 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 372

  51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 373 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 422

  67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 423 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 472

  84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 473 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 522

 101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 523 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 572

 117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 573 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 622

 134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 623 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 672
```

```
    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    673 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 722

    167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    723 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 772

    184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    773 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 822

    201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    823 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 872

    217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    873 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 922

    234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    923 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 972

    251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    973 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1022

    267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
        ||||||||||||||||||||||||||||||
   1023 CATGCTCACAAAAGAACTCTATTTTTATCAC 1053

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAC64043

seq_documentation_block:
ID   AAC64043 standard; DNA; 2538 BP.
XX
AC   AAC64043;
XX
DT   14-FEB-2001  (first entry)
XX
DE   pYNGCas3p10p35 insert encoding human caspase-3 and AcNPV p35 protein.
XX
KW   Human caspase-3; AcNPV p35 protein; BmNPV p10 promoter;
KW   polyhedron promoter; baculovirus vector; apoptosis inhibitor protein;
KW   recominant caspase production; insect cell; ds.
XX
OS   Chimeric - Homo sapiens.
OS   Chimeric - Autographa californica nuclear polyhedrosis virus.
OS   Chimeric - Bombyx mori nuclear polyhedrosis virus.
XX
PN   JP2000253882-A.
XX
PD   19-SEP-2000.
XX
PF   10-MAR-1999;   99JP-0062732.
XX
PR   10-MAR-1999;   99JP-0062732.
XX
PA   (KATA ) KATAKURA IND CO LTD.
XX
DR   WPI; 2000-658811/64.
XX
PT   Baculovirus vector used for the preparation of caspase protein,
PT   comprises a transfer vector plasmid containing a caspase
PT   protein-producing gene -
XX
PS   Example 1; Page 7-8; 10pp; Japanese.
XX
CC   The invention relates to a novel baculovirus vector used for the
CC   preparation of a caspase protein, and containing a recombinant transfer
CC   vector plasmid comprising a gene encoding a caspase and a gene encoding
CC   an apoptosis-inhibiting protein. The invention also encompasses the
CC   preparation of caspase protein, comprising transfecting the novel
CC   baculovirus vector into an insect or an insect cell; breeding the
CC   insect or culturing the cell; and collecting the caspase protein from
CC   the insect or insect cell. The present sequence represents the expression
CC   cassette insert from a transfer vector (pYNGCas3p10p35) produced in an
CC   exemplification of the invention. This insert comprises a mature human
CC   caspase-3 gene under the control of a polyhedron promoter, and an
CC   Autographa californica nuclear polyhedrosis virus (AcNPV) p35 gene under
CC   the control of a Bombyx mori nuclear polyhedrosis virus (BmNPV) p10
CC   promoter.
XX
SQ   Sequence 2538 BP; 865 A; 446 C; 506 G; 721 T; 0 other;


alignment_scores:
            Quality: 1463.00              Length:    277
              Ratio:  5.282                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AAC64043   ..

 Align seg 1/1  to: AAC64043  from: 1  to: 2538

      1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    136 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 185

     17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    186 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 235

     34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    236 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 285

     51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    286 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 335

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    336 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 385

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    386 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 435

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    436 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 485

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    486 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 535

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    536 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 585

    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    586 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 635

    167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    636 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 685

    184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    686 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 735
```

```
     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     736 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 785

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     786 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 835

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     836 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 885

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     886 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 935

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     936 CATGCTCACAAAAGAACTCTATTTTTATCAC 966

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81219

seq_documentation_block:
ID   AAX81219 standard; DNA; 835 BP.
XX
AC   AAX81219;
XX
DT   10-SEP-1999  (first entry)
XX
DE   Nucleotide sequence of caspase-3 (CPP32).
XX
KW   Rev-caspase; cysteine protease; zymogen; caspase; cancer; gene therapy;
KW   autoimmune disease; caspase-mediated apoptosis; neurodegenerative;
KW   tumour cell; myocardial infarction; human; ds.
XX
OS   Homo sapiens.
XX
PN   WO9935277-A2.
XX
PD   15-JUL-1999.
XX
PF   11-JAN-1999;   99WO-US00632.
XX
PR   09-JAN-1998;   98US-0070987.
XX
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES;
XX
DR   WPI; 1999-419353/35.
DR   P-PSDB; AAY21717.
XX
PT   New isolated nucleic acid molecule encoding a rev-caspase  - used
PT   for screening and identifying inhibitors or enhancers for treating
PT   cancer or autoimmune disease
XX
PS   Disclosure; Fig 13A-B; 74pp; English.
XX
CC   The invention relates to nucleic acid molecules encoding rev-caspases.
CC   Rev-caspases are cysteine proteases that specifically cleave proteins
CC   after Asp residues and is expressed as a zymogen, in which the small
CC   subunit is N-terminal to a large subunit. A gene delivery vehicle
CC   comprising a rev-caspase coding sequence is useful for the treatment of
CC   cancer, where the gene delivery vehicle is internalised by tumour cells.
CC   The gene delivery vehicle can also be used to treat autoimmune diseases.
CC   Cells transfected with a rev-caspase expressing vector can be used in
CC   identification of inhibitors or enhancers of caspase-mediated apoptosis.
CC   In vitro translated rev-caspase can be used to identify an inhibitor or
CC   enhancer of caspase processing activity. Caspase inhibitors are useful
CC   for treating neurodegenerative diseases as well as for inhibiting
CC   apoptosis in the heart following myocardial infarction. Sequences
CC   AAX81217 -AAX81226 represent human caspase genes encoding caspase 1-10
CC   gene products (AAY21715-Y21724).
XX
SQ   Sequence 835 BP; 274 A; 147 C; 178 G; 235 T; 1 other;

alignment_scores:
            Quality: 1460.00              Length:    277
              Ratio:  5.271                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.639

alignment_block:
 US-09-613-508-4 x AAX81219   ..

 Align seg 1/1  to: AAX81219  from: 1  to: 835

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 50

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 100

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 150

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 200

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 250

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 300

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 350

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 400

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 450

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 500

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 550

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||:::|||||||||||||||||||||||||||||||
     551 GTCATAAAATACCAGTGGATGCCGACTTCTTGTATGCATACTCCACAGCA 600

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     601 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 650

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     651 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 700

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     701 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 750
```

```
     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     751 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 800

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         |||||||||||||||||||||||||||||
     801 CATGCTCACAAAAGAACTCTATTTTTATCAC 831

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAS10503

seq_documentation_block:
ID   AAS10503 standard; DNA; 2635 BP.
XX
AC   AAS10503;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Human caspase 3 DNA sequence.
XX
KW   Human; caspase 3; apoptosis; hyperproliferative disorder; hepatitis;
KW   viral infection; haematopoietic disorder; autoimmune disorder;
KW   atherosclerosis; neurological disorder; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             225..1058
FT                   /*tag= a
FT                   /product= "Caspase 3"
XX
PN   WO200153310-A1.
XX
PD   26-JUL-2001.
XX
PF   11-JAN-2001; 2001WO-US00888.
XX
PR   18-JAN-2000; 2000US-0484617.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Zhang H,  Cowsert LM;
XX
DR   WPI; 2001-442252/47.
DR   P-PSDB; AAU05394.
XX
PT   New antisense compound to inhibit caspase 3 is useful for treating
PT   hepatitis and atherosclerosis -
XX
PS   Example 13; Page 92-94; 127pp; English.
XX
CC   The present sequence encoding for human caspase 3 is described
CC   in an invention relating to novel antisense oligonucleotides
CC   (AAS10517-AAS10676) and methods of using these compounds for the
CC   modulation of caspase 3 expression. The caspase 3 antisense
CC   oligonucleotides specifically hybridise with and inhibit the
CC   expression of caspase 3. Antisense compounds targeted to caspase 3
CC   are useful to inhibit caspase 3 expression in cells or tissues and
CC   to modulate apoptosis. The caspase 3 antisense oligonucleotides are
CC   useful for treating disorders associated with expression of caspase 3.
CC   Such disorders include hyperproliferative disorders (e.g. cancer),
CC   viral infections (e.g. hepatitis), haematopoietic disorders, autoimmune
CC   disorders, atherosclerosis and neurological disorders (e.g. Alzheimer's
CC   disease).
XX
SQ   Sequence 2635 BP; 815 A; 465 C; 569 G; 786 T; 0 other;


alignment_scores:
           Quality: 1460.00            Length:    277
             Ratio:  5.271               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.639
```

```
alignment_block:
 US-09-613-508-4 x AAS10503   ..

 Align seg 1/1  to: AAS10503  from: 1  to: 2635

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     225 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 274

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     275 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 324

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     325 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 424

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     425 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 474

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     475 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 524

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     525 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 574

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     575 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 624

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     625 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 674

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 724

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     725 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 774

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||:::||||||||||||||||||||||||||||||
     775 GTCATAAAATACCAGTGGATGCCGACTTCTTGTATGCATACTCCACAGCA 824

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     825 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 874

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     875 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 924

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     925 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 974

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     975 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1024

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         |||||||||||||||||||||||||||||
    1025 CATGCTCACAAAAGAACTCTATTTTTATCAC 1055
```

```
seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03910

seq_documentation_block:
ID   AAD03910 standard; cDNA; 2635 BP.
XX
AC   AAD03910;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Human caspase-3 cDNA.
XX
KW   Human; caspase-3; CPP32; yama; apopain;
KW   cysteine protease; apoptosis; caspase expression cassette; metastasis;
KW   tumour; cathepsin B; urokinase; proliferation; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             225..1058
FT                   /*tag= a
FT                   /product= "Human caspase-3"
XX
PN   WO200129232-A2.
XX
PD   26-APR-2001.
XX
PF   19-OCT-2000; 2000WO-US28941.
XX
PR   20-OCT-1999;  99US-0160559.
PR   14-AUG-2000; 2000US-0225564.
XX
PA   (SCIO-) SCIOS INC.
XX
PI   Cordell B,  Li Y;
XX
DR   WPI; 2001-290920/30.
DR   P-PSDB; AAE00600.
XX
PT   Novel fusion polypeptide comprising first and second caspase subunit
PT   separated by cleavage site not associated in nature with caspase
PT   subunit, useful for cloning gene encoding enzymes involved in
PT   proteolytic cleavage  -
XX
PS   Example 1; Fig 5; 116pp; English.
XX
CC   The present sequence is a cDNA encoding human Caspase-3 also known as
CC   CPP32, yama and apopain. Caspases are a family of cysteine proteases,
CC   that participate in the initiation and execution of apoptosis. Caspases
CC   exist as pro-enzymes, activated by cleavage into a large and small
CC   subunit, occurring after specific aspartic acid residues within the
CC   pro-enzyme sequence. The present invention relates to a method for
CC   functional cloning of genes encoding proteins or enzymes involved in
CC   proteolytic cleavage. The invention is based on the use of caspase
CC   expression cassettes comprising the coding sequence of a proteolytic
CC   cleavage site flanked by sequences encoding two caspase subunits. A
CC   fusion polypeptide comprising a first and a second caspase subunit,
CC   separated by a cleavage site not associated in nature, is useful for
CC   cloning gene encoding enzymes involved in proteolytic cleavage. An
CC   expression cassette containing fusion polypeptide is used to identify a
CC   mutant cell line deficient in an enzyme of interest and is also useful
CC   for diagnosis and suppression of proliferation or metastases of a tumour
CC   cell characterised by overexpression of a polypeptide (e.g. Cathepsin B
CC   or urokinase, selectively expressed in the tumour cells). DNA encoding
CC   fusion polypeptide is used in gene therapy.
XX
SQ   Sequence 2635 BP; 815 A; 465 C; 569 G; 786 T; 0 other;


alignment_scores:
           Quality: 1460.00              Length:    277
             Ratio:  5.271                Gaps:      0
Percent Similarity: 100.000   Percent Identity: 99.639
```

```
alignment_block:
 US-09-613-508-4 x AAD03910   ..

 Align seg 1/1  to: AAD03910  from: 1  to: 2635

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     225 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 274

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     275 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 324

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     325 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 424

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     425 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 474

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     475 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 524

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     525 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 574

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     575 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 624

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     625 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 674

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 724

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     725 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 774

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||:::||||||||||||||||||||||||||||||||
     775 GTCATAAAATACCAGTGGATGCCGACTTCTTGTATGCATACTCCACAGCA 824

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     825 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 874

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     875 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 924

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     925 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 974

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     975 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1024

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
    1025 CATGCTCACAAAAGAACTCTATTTTTATCAC 1055
```

```
seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV32608

seq_documentation_block:
ID   AAV32608 standard; cDNA; 1001 BP.
XX
AC   AAV32608;
XX
DT   26-OCT-1998  (first entry)
XX
DE   Mutant human apopain (caspase-3) pro-enzyme cDNA.
XX
KW   Mutant human apopain pro-enzyme; caspase-3; protease; SPA;
KW   scintillation proximity binding assay; diabetes; cancer;
KW   4-phosphono(difluoromethyl) phenylalanine group; osteoporosis;
KW   tyrosine phosphatase; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             64..897
FT                   /*tag= a
FT                   /product= "Mutant apopain C163S pro-enzyme"
FT   misc_feature    148..588
FT                   /*tag= b
FT                   /note= "Encodes the P17 subunit"
FT   mutation        550..552
FT                   /*tag= c
FT                   /note= "Changed from TGC in wild-type to TCC in
FT                    mutant"
FT   misc_feature    589..894
FT                   /*tag= d
FT                   /note= "Encodes the P12 subunit"
XX
PN   WO9820024-A1.
XX
PD   14-MAY-1998.
XX
PF   03-NOV-1997;   97WO-CA00824.
XX
PR   04-NOV-1996;   96US-0030411.
XX
PA   (MERI ) MERCK FROSST CANADA INC.
XX
PI   Desmarais S,  Friesen R,  Zamboni R;
XX
DR   WPI; 1998-348101/30.
DR   P-PSDB; AAW48937.
XX
PT   Peptide(s) useful in binding assays for tyrosine phosphatases or
PT   cysteine proteases - contain two or more 4-phosphono(difluoromethyl)
PT   phenylalanine groups to improve binding affinity
XX
PS   Disclosure; Pages 38-39; 59pp; English.
XX
CC   The present sequence represents a mutated human apopain (caspase-3)
CC   pro-enzyme cDNA.  The cDNA encodes a mutant human apopain C163S
CC   pro-enzyme.  The invention provides a method for use in a
CC   scintillation proximity binding assay (SPA) for proteases and
CC   phosphatases.  The method involves using mutated proteases and
CC   phosphatases whereby the catalytic cysteine residue of the enzymes
CC   are replaced with a serine or alanine residue to correct the problem of
CC   interference in SPA from extraneous oxidising and alkylating agents.
CC   The mutation affects the catalytic properties of the enzyme but does
CC   not affect their binding properties.  The invention claims for new
CC   ligands for use in SPA which have increased binding affinity for a
CC   tyrosine phosphatase or cysteine protease.  The ligands contain at least
CC   two 4-phosphono(difluoromethyl) phenylalanine groups which increase
CC   binding affinity of the ligand to its respective enzyme.  The assay can
CC   be used to determine the ability of new ligands and compound mixtures to
CC   competitively bind with an enzyme.  The method is claimed to allow
CC   a better usage of SPA in the discovery of compounds for the treatment
CC   and study of diseases, e.g. diabetes, cancer and osteoporosis.
XX
SQ   Sequence 1001 BP; 310 A; 172 C; 226 G; 293 T; 0 other;


alignment_scores:
            Quality: 1453.00              Length:    277
              Ratio:  5.264                Gaps:      0
 Percent Similarity:  99.639   Percent Identity: 99.639

alignment_block:
 US-09-613-508-4 x AAV32608   ..

 Align seg 1/1  to: AAV32608  from: 1  to: 1001

      1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

     17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

     34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

     51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        ||||||||||||||||||||||||||||||||||  ||||||||||||||
    514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTCCCGTGGTACAGA 563

    167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

    184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

    201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

    217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

    234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813
```

```
     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV32615

seq_documentation_block:
ID   AAV32615 standard; cDNA; 1001 BP.
XX
AC   AAV32615;
XX
DT   26-OCT-1998  (first entry)
XX
DE   Mutant human apopain (caspase-3) pro-enzyme cDNA.
XX
KW   Mutant human apopain pro-enzyme; caspase-3; protease; SPA;
KW   scintillation proximity binding assay; diabetes; cancer;
KW   4-phosphono(difluoromethyl) phenylalanine group; osteoporosis;
KW   tyrosine phosphatase; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             64..897
FT                   /*tag= a
FT                   /product= "Mutant apopain C163S pro-enzyme"
FT   misc_feature    148..588
FT                   /*tag= b
FT                   /note= "Encodes the P17 subunit"
FT   mutation        550..552
FT                   /*tag= c
FT                   /note= "Changed from TGC in wild-type to TCC in
FT                   mutant"
FT   misc_feature    589..894
FT                   /*tag= d
FT                   /note= "Encodes the P12 subunit"
XX
PN   WO9820156-A1.
XX
PD   14-MAY-1998.
XX
PF   03-NOV-1997;   97WO-CA00825.
XX
PR   04-NOV-1996;   96US-0030408.
XX
PA   (MERI ) MERCK FROSST CANADA INC.
XX
PI   Desmarais S,  Ford-Hutchinson A,  Friesen R,  Gresser M;
PI   Kennedy B,  Nicholson D,  Ramachandran C,  Skorey K;
XX
DR   WPI; 1998-348103/30.
XX
PT   Enzyme binding assay for detection of useful compounds - uses mutant
PT   form of wild-type enzyme, in which serine replaces cysteine at
PT   active site, to reduce interference from oxidising/alkylating agents
XX
PS   Disclosure; Pages 38-39; 63pp; English.
XX
CC   The present sequence represents a mutated human apopain (caspase-3)
CC   pro-enzyme cDNA.  The cDNA encodes a mutant human apopain C163S
CC   pro-enzyme.  The invention claims for a method for use in a
CC   scintillation proximity binding assay (SPA) for proteases and
CC   phosphatases.  The method involves using mutated proteases and
CC   phosphatases whereby the catalytic cysteine residue of the enzymes
CC   are replaced with a serine or alanine residue to correct the problem of
CC   interference in SPA from extraneous oxidising and alkylating agents.
CC   The mutation affects the catalytic properties of the enzyme but does
CC   not affect their binding properties.  The invention provides new
CC   ligands for use in SPA which have increased binding affinity for a
CC   tyrosine phosphatase or cysteine protease.  The ligands contain at least
CC   two 4-phosphono(difluoromethyl) phenylalanine groups which increase
CC   binding affinity of the ligand to its respective enzyme.  The assay can
CC   be used to determine the ability of new ligands and compound mixtures to
CC   competitively bind with an enzyme.  The method is claimed to allow
CC   a better usage of SPA in the discovery of compounds for the treatment
CC   and study of diseases, e.g. diabetes, cancer and osteoporosis.
XX
SQ   Sequence 1001 BP; 310 A; 172 C; 226 G; 293 T; 0 other;


alignment_scores:
           Quality: 1453.00              Length:    277
             Ratio:  5.264                Gaps:      0
 Percent Similarity:  99.639   Percent Identity: 99.639

alignment_block:
 US-09-613-508-4 x AAV32615   ..

 Align seg 1/1  to: AAV32615  from: 1  to: 1001

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||  ||||||||||||
     514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTCCCGTGGTACAGA 563

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
      664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

      217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

      234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

      251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

      267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
          ||||||||||||||||||||||||||||||
      864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13078

seq_documentation_block:
ID   AAD13078 standard; DNA; 6436 BP.
XX
AC   AAD13078;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-1XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-1XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain -
XX
PS   Example 4; Page 133-134; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-1XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domain-Mammalia caspase-3 fusion protein.
XX
SQ   Sequence 6436 BP; 1577 A; 1606 C; 1607 G; 1646 T; 0 other;


alignment_scores:
           Quality: 1453.00              Length:    275
             Ratio:  5.284                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AAD13078   ..

 Align seg 1/1  to: AAD13078  from: 1  to: 6436

        3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1144 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAA 1193

       19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1194 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1243

       36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1244 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 1293

       53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1294 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 1343

       69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1344 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 1393

       86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1394 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1443

      103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1444 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 1493

      119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1494 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 1543

      136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1544 TGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 1593

      153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1594 GGAAAACCCAAACTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 1643

      169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1644 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 1693

      186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1694 AAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 1743

      203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1744 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 1793

      219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1794 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 1843

      236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||
  1844 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 1893

   253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1894 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 1943

   269 uThrLysGluLeuTyrPheTyrHis 277
       ||||||||||||||||||||||||
  1944 CACAAAAGAACTCTATTTTTATCAC 1968

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13079

seq_documentation_block:
ID   AAD13079 standard; DNA; 6703 BP.
XX
AC   AAD13079;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-2XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-2XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain  -
XX
PS   Example 4; Page 134-135; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-2XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domains-Mammalia caspase-3
CC   fusion protein.
XX
SQ   Sequence 6703 BP; 1653 A; 1677 C; 1674 G; 1699 T; 0 other;

alignment_scores:
            Quality: 1453.00              Length:    275
              Ratio:  5.284                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AAD13079   ..

 Align seg 1/1  to: AAD13079  from: 1  to: 6703

      3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1411 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAA 1460

     19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1461 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1510

     36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1511 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 1560

     53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1561 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 1610

     69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1611 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 1660

     86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1661 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1710

    103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1711 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 1760

    119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1761 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 1810

    136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1811 TGAAAAAAATAACAAACTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 1860

    153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1861 GGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 1910

    169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1911 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 1960

    186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1961 AAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2010

    203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   2011 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 2060

    219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   2061 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 2110

    236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   2111 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 2160

    253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   2161 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 2210
```

```
     269 uThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||
    2211 CACAAAAGAACTCTATTTTTATCAC 2235

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13080

seq_documentation_block:
ID   AAD13080 standard; DNA; 6946 BP.
XX
AC   AAD13080;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-3XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-3XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain  -
XX
PS   Example 4; Page 136-137; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-3XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domains-Mammalia caspase-3
CC   fusion protein.
XX
SQ   Sequence 6946 BP; 1723 A; 1742 C; 1735 G; 1746 T; 0 other;


alignment_scores:
           Quality: 1453.00              Length:    275
             Ratio:  5.284                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:

US-09-613-508-4 x AAD13080   ..

Align seg 1/1  to: AAD13080  from: 1  to: 6946

       3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1654 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAA 1703

      19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1704 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1753

      36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1754 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 1803

      53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1804 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 1853

      69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1854 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 1903

      86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1904 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1953

     103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1954 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 2003

     119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2004 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 2053

     136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2054 TGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 2103

     153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2104 GGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 2153

     169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2154 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 2203

     186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2204 AAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2253

     203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2254 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 2303

     219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2304 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 2353

     236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2354 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 2403

     253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2404 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 2453

     269 uThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||
    2454 CACAAAAGAACTCTATTTTTATCAC 2478

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13081
```

```
seq_documentation_block:
ID   AAD13081 standard; DNA; 7189 BP.
XX
AC   AAD13081;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-4XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-4XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain  -
XX
PS   Example 4; Page 137-138; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-4XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domains-Mammalia caspase-3
CC   fusion protein.
XX
SQ   Sequence 7189 BP; 1793 A; 1807 C; 1796 G; 1793 T; 0 other;


alignment_scores:
            Quality: 1453.00              Length:    275
              Ratio:  5.284                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AAD13081   ..

 Align seg 1/1  to: AAD13081  from: 1  to: 7189

       3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1897 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAA 1946

      19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1947 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1996

      36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1997 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 2046

      53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2047 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 2096

      69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2097 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 2146

      86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2147 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 2196

     103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2197 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 2246

     119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2247 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 2296

     136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2297 TGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 2346

     153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2347 GGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 2396

     169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2397 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 2446

     186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2447 AAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2496

     203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2497 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 2546

     219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2547 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 2596

     236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2597 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 2646

     253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2647 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 2696

     269 uThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||
    2697 CACAAAAGAACTCTATTTTTATCAC 2721

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13084

seq_documentation_block:
ID   AAD13084 standard; DNA; 7247 BP.
XX
AC   AAD13084;
XX
```

```
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-Ub-Met-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-Ub-Met-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain  -
XX
PS   Example 4; Page 140-141; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-Ub-Met-C3 construct comprising wildtype multimerised
CC   human ubiquitin destabilisation domain-Mammalia caspase-3 fusion protein.
XX
SQ   Sequence 7247 BP; 1794 A; 1796 C; 1826 G; 1831 T; 0 other;


alignment_scores:
            Quality: 1442.00              Length:    275
              Ratio:  5.263                Gaps:      0
 Percent Similarity:  99.636   Percent Identity: 99.273

alignment_block:
 US-09-613-508-4 x AAD13084   ..

 Align seg 1/1  to: AAD13084  from: 1  to: 7247

       3 AsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGluProLy 19
         ||||||||||||||||||   ||||||||||||||||||||||||||||||
    1150 AACACTGAAAACTCAGTGGCCTCAAAATCCATTAAAAATTTGGAACCAAA 1199

      19 sIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnS 36
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1200 GATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACA 1249

      36 erTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsn 52
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1250 GTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAAT 1299

      53 LysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspVa 69
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1300 AAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGT 1349

      69 lAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValA 86
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1350 CGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCA 1399

      86 rgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAsp 102
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1400 GGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1449

      103 ValSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLe 119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1450 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCT 1499

      119 uSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspL 136
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1500 GAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACC 1549

      136 euLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThr 152
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1550 TGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACT 1599

      153 GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAs 169
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1600 GGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGA 1649

      169 pCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysHisL 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1650 CTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATA 1699

      186 ysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGly 202
          ||||||||||||:::||||||||||||||||||||||||||||||||||||
    1700 AAATACCAGTGGATGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 1749

      203 TyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLe 219
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1750 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCT 1799

      219 uCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleL 236
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1800 TTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTC 1849

      236 euThrArgValAsnArgLysValAlaThrGluPheGluSerPheSerPhe 252
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1850 TTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTT 1899

      253 AspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerMetLe 269
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1900 GACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCT 1949

      269 uThrLysGluLeuTyrPheTyrHis 277
          ||||||||||||||||||||||||
    1950 CACAAAAGAACTCTATTTTTATCAC 1974
```

OM of: US-09-613-508-4 to: Issued_Patents_NA:* out_format : pfs

Date: Aug 15, 2002 8:47 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09613508/runat_14082002_112647_7244/app_query.fasta_1.701
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09613508_@CGN1_1_54 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-613-508-4
Query length: 277
Database: Issued_Patents_NA:*
Database sequences: 383533
Database length: 122816752
Search time (sec): 63.980000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-308-5	+	1463.00	3342.20	1.8e-178	1001	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-313-5	+	1463.00	3342.20	1.8e-178	1001	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-069-138-5	+	1463.00	3342.20	1.8e-178	1001	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-591-605-1	+	1463.00	3340.44	2.2e-178	1158	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-3	+	1463.00	3340.43	2.2e-178	1159	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-3	+	1460.00	3323.56	2.0e-177	2635	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-308-9	+	1453.00	3319.13	3.4e-177	1001	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-313-9	+	1453.00	3319.13	3.4e-177	1001	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-069-138-9	+	1453.00	3319.13	3.4e-177	1001	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-890-542A-1	+	1304.00	2973.43	6.2e-158	1177	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-890-542A-3	+	1304.00	2973.43	6.2e-158	1177	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-10	+	1293.00	2946.88	1.9e-156	1297	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-1	+	712.00	1605.86	9.3e-82	1371	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-446-925-4	+	517.50	1155.71	1.1e-56	1545	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-146-331-4	+	517.50	1155.71	1.1e-56	1545	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-896-885-4	+	517.50	1155.71	1.1e-56	1545	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-375-256-4	+	517.50	1155.71	1.1e-56	1545	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-522-813-1	+	516.50	1156.64	9.8e-57	1183	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-773-608A-1	+	470.00	1039.56	3.2e-50	2658	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-487-445-10	+	435.00	965.83	4.1e-46	1489	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-360-017-2	+	423.00	944.65	6.3e-45	870	S
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-807-200-1	+	422.00	937.21	1.6e-44	1330	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-001-777-1	+	422.00	937.21	1.6e-44	1330	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-807-200-3	+	421.00	935.31	2.1e-44	1286	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-001-777-3	+	421.00	935.31	2.1e-44	1286	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-618-408B-3	+	421.00	934.68	2.3e-44	1355	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-665-220-3	+	421.00	930.69	3.8e-44	1883	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-487-445-3	+	421.00	930.69	3.8e-44	1883	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-291-692-3	+	421.00	930.69	3.8e-44	1883	
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US96-10521-17	+	421.00	926.70	6.3e-44	2619	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-718-964-1	+	421.00	926.56	6.4e-44	2649	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-059-964A-1	+	421.00	926.56	6.4e-44	2649	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-842-341-1	+	421.00	926.56	6.4e-44	2649	
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US96-10521-14	+	421.00	925.52	7.3e-44	2887	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-807-200-11	+	419.00	929.33	4.5e-44	1440	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-001-777-11	+	419.00	929.33	4.5e-44	1440	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-446-925-6	+	418.50	929.29	4.5e-44	1313	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-146-331-6	+	418.50	929.29	4.5e-44	1313	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-896-885-6	+	418.50	929.29	4.5e-44	1313	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-375-256-6	+	418.50	929.29	4.5e-44	1313	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-852-936C-2	+	418.00	925.91	6.9e-44	1578	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-300-328-2	+	418.00	925.91	6.9e-44	1578	
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US96-10521-33	+	417.00	924.69	8.1e-44		
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-257-218-1	+	410.00	910.27	5.2e-43	1251	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-311-760-1	+	410.00	910.27	5.2e-43	1251	

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-308-5

seq_documentation_block:
; Sequence 5, Application US/08964308
; Patent No. 6066715
; GENERAL INFORMATION:
; APPLICANT: DESMARAIS, SYLVIE
; APPLICANT: FRIESEN, RICHARD
; APPLICANT: ZAMBONI, ROBERT
; TITLE OF INVENTION: NEW LIGANDS FOR PHOSPHATASE
; TITLE OF INVENTION: BINDING ASSAY
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBERT J. NORTH - MERCK & CO., INC.
; STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Diskette
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/964,308
; FILING DATE: 04-NOV-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: NORTH, ROBERT J
; REGISTRATION NUMBER: 27,366
; REFERENCE/DOCKET NUMBER: 19840 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-7262
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-964-308-5

alignment_scores:
Quality: 1463.00 Length: 277
Ratio: 5.282 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-613-508-4 x US-08-964-308-5 ..

Align seg 1/1 to: US-08-964-308-5 from: 1 to: 1001

```
       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213
```

```
     51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 563

    167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

    184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

    201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

    217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

    234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

    251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

    267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
        |||||||||||||||||||||||||||||
    864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-313-5

seq_documentation_block:
; Sequence 5, Application US/08964313
; Patent No. 6114132
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  GRESSER, MICHAEL
;    APPLICANT:  KENNEDY, BRIAN
;    APPLICANT:  NICHOLSON, DONALD
;    APPLICANT:  RAMACHANDRAN, CHIDAMBARAN
;    APPLICANT:  SKOREY, KATHRYN
;    APPLICANT:  FORD-HUTCHINSON, ANTHONY
;    TITLE OF INVENTION:  PHOSPHATASE BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,313
;      FILING DATE:  04-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/030,408
;      FILING DATE:  04-NOV-1996
;      APPLICATION NUMBER:  PCT/CA97/00825
;      FILING DATE:  03 -NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19824Y
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-313-5


alignment_scores:
           Quality: 1463.00              Length:    277
             Ratio:  5.282                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x US-08-964-313-5   ..

 Align seg 1/1  to: US-08-964-313-5  from: 1  to: 1001

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363
```

```
     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 563

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-069-138-5

seq_documentation_block:
; Sequence 5, Application US/09069138
; Patent No. 6348572
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  DUFRESNE, CLAUDE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  LEBLANC, YVES
;    APPLICANT:  ROY, PATRICK
;    APPLICANT:  YOUNG, ROBERT N.
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/069,138
;      FILING DATE:  29-APR-1998
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19840YIA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-069-138-5


alignment_scores:
           Quality: 1463.00                Length:    277
             Ratio:  5.282                   Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x US-09-069-138-5   ..

 Align seg 1/1  to: US-09-069-138-5  from: 1  to: 1001

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 563

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-591-605-1

seq_documentation_block:
; Sequence 1, Application US/08591605
; Patent No. 6060238
;  GENERAL INFORMATION:
;    APPLICANT:  Dixit, Vishva M.
;    TITLE OF INVENTION:  METHOD AND COMPOSITION FOR REGULATING
;    TITLE OF INVENTION:  APOPTOSIS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 PAGE MILL ROAD
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/591,605
;      FILING DATE:  09-FEB-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KONSKI, ANTOINETTE F.
;      REGISTRATION NUMBER:  34,202
;      REFERENCE/DOCKET NUMBER:  20344-21036.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  706141 MRSNFOERS SFO
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1158 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  224..1054
;      OTHER INFORMATION:  /product= "Yama peptide sequence"
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  704..718
;      OTHER INFORMATION:  /note= "region coding for
;      OTHER INFORMATION:  pentapeptide, wherein cysteine is catalytic"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "atg")
;      OTHER INFORMATION:  /note= "Mutation of catalytic cysteine to methionine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gcc")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gcg")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gct")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gca")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
US-08-591-605-1


alignment_scores:
            Quality: 1463.00              Length:    277
              Ratio:  5.282                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x US-08-591-605-1   ..

 Align seg 1/1  to: US-08-591-605-1  from: 1  to: 1158

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     224 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 273

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     274 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 323

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     324 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 373

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     374 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 423

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     424 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 473

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     474 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 523

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     524 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 573

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     574 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 623

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     624 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 673

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     674 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 723

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     724 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 773

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     774 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 823

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     824 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 873

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     874 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 923

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     924 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 973

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     974 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1023

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         |||||||||||||||||||||||||||||||
    1024 CATGCTCACAAAAGAACTCTATTTTTATCAC 1054

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-3

seq_documentation_block:
; Sequence 3, Application US/08462969B
; Patent No. 6087150
;  GENERAL INFORMATION:
;    APPLICANT:  He, Wei-Wu et al.
;    TITLE OF INVENTION:  Interleukin-1 Beta Converting Enzyme
;    TITLE OF INVENTION:  Like Apoptosis Protease 3 and 4
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Ave.
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,969B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/334,251
;      FILING DATE:  11-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF140P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1159 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)

US-08-462-969B-3

alignment_scores:
           Quality: 1463.00              Length:    277
             Ratio:  5.282                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x US-08-462-969B-3   ..

 Align seg 1/1  to: US-08-462-969B-3  from: 1  to: 1159

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     223 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 272

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     273 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 322

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     323 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 372

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     373 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 422

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     423 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 472

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     473 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 522

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     523 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 572

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     573 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 622

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     623 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 672

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     673 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 722

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     723 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 772

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     773 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 822

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     823 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 872

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     873 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 922

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     923 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 972
```

```
     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     973 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1022

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         |||||||||||||||||||||||||||||
    1023 CATGCTCACAAAAGAACTCTATTTTTATCAC 1053

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-3

seq_documentation_block:
; Sequence 3, Application US/09484617
; Patent No. 6303374
; GENERAL INFORMATION:
;  APPLICANT: Hong Zhang
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF CASPASE 3 EXPRESSION
;  FILE REFERENCE: RTS-0103
;  CURRENT APPLICATION NUMBER: US/09/484,617
;  CURRENT FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 176
; SEQ ID NO 3
;   LENGTH: 2635
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (225)...(1058)
US-09-484-617-3


alignment_scores:
            Quality: 1460.00              Length:    277
              Ratio:  5.271                Gaps:      0
 Percent Similarity: 100.000    Percent Identity: 99.639

alignment_block:
 US-09-613-508-4 x US-09-484-617-3   ..

 Align seg 1/1  to: US-09-484-617-3  from: 1  to: 2635

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     225 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 274

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     275 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 324

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     325 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 374

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     375 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 424

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     425 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 474

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     475 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 524

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     525 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 574

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     575 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 624

      134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      625 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 674

      151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      675 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 724

      167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      725 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 774

      184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
          ||||||||||||||||:::||||||||||||||||||||||||||||||||
      775 GTCATAAAATACCAGTGGATGCCGACTTCTTGTATGCATACTCCACAGCA 824

      201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      825 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 874

      217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      875 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 924

      234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      925 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 974

      251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      975 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 1024

      267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
          |||||||||||||||||||||||||||||
     1025 CATGCTCACAAAAGAACTCTATTTTTATCAC 1055

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-308-9

seq_documentation_block:
; Sequence 9, Application US/08964308
; Patent No. 6066715
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBERT J. NORTH - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,308
;      FILING DATE:  04-NOV-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  NORTH, ROBERT J
;      REGISTRATION NUMBER:  27,366
;      REFERENCE/DOCKET NUMBER:  19840 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-7262
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-308-9


alignment_scores:
            Quality: 1453.00              Length:    277
              Ratio:  5.264                Gaps:      0
 Percent Similarity:  99.639   Percent Identity: 99.639

alignment_block:
 US-09-613-508-4 x US-08-964-308-9   ..

 Align seg 1/1  to: US-08-964-308-9  from: 1  to: 1001

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||  ||||||||||||||||
     514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTCCCGTGGTACAGA 563

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-313-9

seq_documentation_block:
; Sequence 9, Application US/08964313
; Patent No. 6114132
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  GRESSER, MICHAEL
;    APPLICANT:  KENNEDY, BRIAN
;    APPLICANT:  NICHOLSON, DONALD
;    APPLICANT:  RAMACHANDRAN, CHIDAMBARAN
;    APPLICANT:  SKOREY, KATHRYN
;    APPLICANT:  FORD-HUTCHINSON, ANTHONY
;    TITLE OF INVENTION:  PHOSPHATASE BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,313
;      FILING DATE:  04-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/030,408
;      FILING DATE:  04-NOV-1996
;      APPLICATION NUMBER:  PCT/CA97/00825
;      FILING DATE:  03 -NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19824Y
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-313-9


alignment_scores:
            Quality: 1453.00              Length:    277
              Ratio:  5.264                Gaps:      0
 Percent Similarity:  99.639   Percent Identity: 99.639
```

```
alignment_block:
 US-09-613-508-4 x US-08-964-313-9   ..

 Align seq 1/1  to: US-08-964-313-9  from: 1  to: 1001

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||  ||||||||||||||||
     514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTCCCGTGGTACAGA 563

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894
```

```
seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-069-138-9

seq_documentation_block:
; Sequence 9, Application US/09069138
; Patent No. 6348572
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  DUFRESNE, CLAUDE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  LEBLANC, YVES
;    APPLICANT:  ROY, PATRICK
;    APPLICANT:  YOUNG, ROBERT N.
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/069,138
;      FILING DATE:  29-APR-1998
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19840YIA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-069-138-9


alignment_scores:
           Quality: 1453.00              Length:    277
             Ratio:  5.264                Gaps:      0
 Percent Similarity:  99.639   Percent Identity: 99.639

alignment_block:
 US-09-613-508-4 x US-09-069-138-9   ..

 Align seq 1/1  to: US-09-069-138-9  from: 1  to: 1001

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      64 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 113

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     114 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 163

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     164 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 213
```

```
     51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    214 AATAATAAGAATTTTCATAAGAGCACTGGAATGACATCTCGGTCTGGTAC 263

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    264 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 313

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    314 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 363

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    364 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 413

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    414 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 463

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    464 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 513

    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        |||||||||||||||||||||||||||||||||   ||||||||||||
    514 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTCCCGTGGTACAGA 563

    167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    564 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 613

    184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    614 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 663

    201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    664 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 713

    217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    714 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 763

    234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    764 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 813

    251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    814 TCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTC 863

    267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
        |||||||||||||||||||||||||||||
    864 CATGCTCACAAAAGAACTCTATTTTTATCAC 894

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-890-542A-1

seq_documentation_block:
; Sequence 1, Application US/08890542A
; Patent No. 5840509
;  GENERAL INFORMATION:
;    APPLICANT:  Ni, Binhui
;    APPLICANT:  Paul, Steven M
;    APPLICANT:  Wu, Xin
;    TITLE OF INVENTION:  PROTEASE AND RELATED NUCLEIC ACID
;    TITLE OF INVENTION:  COMPOUNDS
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  United States of America
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/890,542A
;      FILING DATE:  09-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gaylo, Paul J.
;      REGISTRATION NUMBER:  36,808
;      REFERENCE/DOCKET NUMBER:  X-10704
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (317) 276-0756
;      TELEFAX:  (317) 276-3861
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1177 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  76..906
US-08-890-542A-1


alignment_scores:
            Quality: 1304.00              Length:    277
              Ratio:  4.902                Gaps:      0
 Percent Similarity:  96.029   Percent Identity: 88.448

alignment_block:
 US-09-613-508-4 x US-08-890-542A-1   ..

 Align seg 1/1  to: US-08-890-542A-1  from: 1  to: 1177

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         |||:::|||:::|||:::|||||||||||||||||||||||:::|||:::||
      76 ATGGACAACAACGAAACCTCCGTGGATTCAAAATCCATTAATAATTTTGA 125

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         |   |||   |||||||||||:::|||||||||||||||||   ||||
     126 AACAAAGACTATCCATGGAAGCAAGTCGATGGACTCTGGAATATATCTGG 175

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||:::|||||||||||||||||||||||||||||||||||||||||||||
     176 ACAGCAGTTACAAAATGGATTACCCTGAAATGGGCTTGTGTATAATAATT 225

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         |||||||||||||||||||||||||||||:::::|||:::|||||
     226 AATAATAAGAACTTCCATAAAAGCACTGGAATGTCAGCTCGCAATGGTAC 275

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         |||||||||||||||||||||||||||||||      ||||||||||
     276 CGATGTCGATGCAGCTAACCTCAGAGAGACATTCATGGCCCTGAAATACG 325

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         |||||||||||||||||||||||||||||||||||:::|||||||||
     326 AAGTCAGGAATAAAAATGACCTTACTCGTGAAGAAATTATGGAATTGATG 375

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
           :::|||||||||||||||||||||||||||||||||||||||||||||
     376 GATAGTGTTTCTAAGGAAGATCACAGCAAAAGGAGCAGTTTTGTGTGTGT 425

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         |:::|||||||||||||:::||||||:::|||||||||||||||||||||
     426 GATTCTAAGTCATGGAGATGAAGGAGTAATTTTTGGAACGAACGGACCTG 475
```

```
     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||:::|||:::|||||||||||||||   |||||||||
     476 TGGACCTGAAAAAACTAACTAGTTTCTTCAGAGGCGACTACTGCCGGAGT 525

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     526 CTGACTGGAAAGCCGAAACTCTTCATCATTCAGGCCTGCCGAGGTACAGA 575

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||:::|||||||||||||||
     576 GCTGGACTGCGGTATTGAGACAGACAGTGGAACTGACGATGATATGGCAT 625

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||:::|||||||||||||   |||||||||||||||||||||||||||||
     626 GCCAGAAGATACCAGTGGGGGCCGACTTCCTGTATGCTTACTCTACCGCA 675

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||:::|||||||||||||||||||||
     676 CCCGGTTACTATTCCTGGAGAAATTCAAGGGACGGGTCATGGTTCATCCA 725

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||   ||||||   ||||||||||||||||||
     726 GTCACTTTGCGCCATGCTGAAACTGTACGCGCACAAGCTGGAATTCATGC 775

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||   ||||||||||||||||
     776 ACATCCTCACTCGTGTTAACCGGAAGGTGGCCATGGAATTTGAGTCCTTC 825

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         |||:::||||||||||||||||||||||||||||||||||||||||||||
     826 TCCCTGGACGCCACTTTCCACGCAAAGAAACAGATCCCGTGTATTGTGTC 875

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     876 AATGCTCACAAAAGAACTGTACTTTTATCAC 906

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-890-542A-3

seq_documentation_block:
; Sequence 3, Application US/08890542A
; Patent No. 5840509
;  GENERAL INFORMATION:
;    APPLICANT:  Ni, Binhui
;    APPLICANT:  Paul, Steven M
;    APPLICANT:  Wu, Xin
;    TITLE OF INVENTION:  PROTEASE AND RELATED NUCLEIC ACID
;    TITLE OF INVENTION:  COMPOUNDS
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  United States of America
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/890,542A
;      FILING DATE:  09-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gaylo, Paul J.
;      REGISTRATION NUMBER:  36,808
;      REFERENCE/DOCKET NUMBER:  X-10704
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (317) 276-0756
;      TELEFAX:  (317) 276-3861
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1177 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  mRNA
US-08-890-542A-3


alignment_scores:
            Quality: 1304.00              Length:    277
              Ratio:  4.902                Gaps:      0
 Percent Similarity:  96.029   Percent Identity: 88.448

alignment_block:
 US-09-613-508-4 x US-08-890-542A-3   ..

 Align seg 1/1  to: US-08-890-542A-3  from: 1  to: 1177

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         |||:::|||:::|||:::|||||||||||||||||||||||:::|||:::||
      76 AUGGACAACAACGAAACCUCCGUGGAUUCAAAAUCCAUUAAUAAUUUUGA 125

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         |   |||   ||||||||||||:::|||||||||||||||||   ||||
     126 AACAAAGACUAUCCAUGGAAGCAAGUCGAUGGACUCUGGAAUAUAUCUGG 175

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||:::|||||||||||||||||||||||||||||||||||||||||||||
     176 ACAGCAGUUACAAAAUGGAUUACCCUGAAAUGGGCUUGUGUAUAAUAAUU 225

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         |||||||||||||||||||||||||||||||:::::::|||:::||||||
     226 AAUAAUAAGAACUUCCAUAAAAGCACUGGAAUGUCAGCUCGCAAUGGUAC 275

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         |||||||||||||||||||||||||||||||||      |||||||||
     276 CGAUGUCGAUGCAGCUAACCUCAGAGAGACAUUCAUGGCCCUGAAAUACG 325

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||:::|||||||||||
     326 AAGUCAGGAAUAAAAAUGACCUUACUCGUGAAGAAAUUAUGGAAUUGAUG 375

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         :::||||||||||||||||||||||||||||||||||||||||||||||||
     376 GAUAGUGUUUCUAAGGAAGAUCACAGCAAAAGGAGCAGUUUUGUGUGUGU 425

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         |:::|||||||||||||:::|||||||:::|||||||||||||||||||||
     426 GAUUCUAAGUCAUGGAGAUGAAGGAGUAAUUUUUGGAACGAACGGACCUG 475

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||:::|||:::|||||||||||||||   |||||||||
     476 UGGACCUGAAAAAACUAACUAGUUUCUUCAGAGGCGACUACUGCCGGAGU 525

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     526 CUGACUGGAAAGCCGAAACUCUUCAUCAUUCAGGCCUGCCGAGGUACAGA 575

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||:::|||||||||||||||
     576 GCUGGACUGCGGUAUUGAGACAGACAGUGGAACUGACGAUGAUAUGGCAU 625

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||:::|||||||||||||   |||||||||||||||||||||||||||||
     626 GCCAGAAGAUACCAGUGGGGGCCGACUUCCUGUAUGCUUACUCUACCGCA 675

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||:::|||||||||||||||||||||
     676 CCCGGUUACUAUUCCUGGAGAAAUUCAAGGGACGGGUCAUGGUUCAUCCA 725

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||   ||||||   ||||||||||||||||||
     726 GUCACUUUGCGCCAUGCUGAAACUGUACGCGCACAAGCUGGAAUUCAUGC 775
```

```
     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||   ||||||||||||||||
     776 ACAUCCUCACUCGUGUUAACCGGAAGGUGGCCAUGGAAUUUGAGUCCUUC 825

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         |||:::|||||||||||||||||||||||||||||||||||||||||||||
     826 UCCCUGGACGCCACUUUCCACGCAAAGAAACAGAUCCCGUGUAUUGUGUC 875

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     876 AAUGCUCACAAAAGAACUGUACUUUUAUCAC 906

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-10

seq_documentation_block:
; Sequence 10, Application US/09484617
; Patent No. 6303374
; GENERAL INFORMATION:
;  APPLICANT: Hong Zhang
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF CASPASE 3 EXPRESSION
;  FILE REFERENCE: RTS-0103
;  CURRENT APPLICATION NUMBER: US/09/484,617
;  CURRENT FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 176
; SEQ ID NO 10
;   LENGTH: 1297
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (68)...(901)
US-09-484-617-10


alignment_scores:
            Quality: 1293.00              Length:    277
              Ratio:  4.807                Gaps:      0
 Percent Similarity:  97.112   Percent Identity: 86.643

alignment_block:
 US-09-613-508-4 x US-09-484-617-10   ..

 Align seg 1/1  to: US-09-484-617-10  from: 1  to: 1297

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||:::::::::|||||||||||||||||||||:::|||:::||
      68 ATGGAGAACAACAAAACCTCAGTGGATTCAAAATCCATTAATAATTTTGA 117

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         |   |||   ||||||||||||:::|||:::||||||||||||   ||||
     118 AGTAAAGACCATACATGGGAGCAAGTCAGTGGACTCTGGGATCTATCTGG 167

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||:::|||||||||||||||||||||||||||:::|||||||||||||
     168 ACAGTAGTTACAAAATGGATTATCCTGAAATGGGCATATGCATAATAATT 217

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||:::|||||||||||||||
     218 AATAATAAGAACTTCCATAAGAGCACTGGAATGTCATCTCGCTCTGGTAC 267

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||   :::||||||||||:
     268 GGATGTGGACGCAGCCAACCTCAGAGAGACATTCATGGGCCTGAAATACC 317

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ::|||||||||||||||||||||||||||||:::|||:::||||||||||
     318 AAGTCAGGAATAAAAATGATCTTACTCGTGAAGACATTTTGGAATTAATG 367

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
           :::|||||||||||||||||||||||||||||||||||||||||||||
     368 GATAGTGTTTCTAAGGAAGATCATAGCAAAAGGAGCAGCTTTGTGTGTGT 417

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         |:::|||||||||||||:::|||||||:::|||:::|||||||||||||||
     418 GATTCTAAGCCATGGTGATGAAGGGGTCATTTATGGGACAAATGGGCCTG 467

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||:::|||||||||:::|||:::|||||||||||||||   |||||||||
     468 TTGAACTGAAAAAGTTGACTAGCTTCTTCAGAGGCGACTACTGCCGGAGT 517

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     518 CTGACTGGAAAGCCGAAACTCTTCATCATTCAGGCCTGCCGGGGTACGGA 567

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||:::|||:::::::|||||
     568 GCTGGACTGTGGCATTGAGACAGACAGTGGGACTGATGAGGAGATGGCTT 617

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||:::|||||||||||||||||||||||||||||||||||||||||||||
     618 GCCAGAAGATACCGGTGGAGGCTGACTTCCTGTATGCTTACTCTACAGCA 667

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     668 CCTGGTTACTATTCCTGGAGAAATTCAAAGGACGGGTCGTGGTTCATCCA 717

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         |||||||||:::|||||||||   ||||||   ||||||||||||||||
     718 GTCCCTTTGCAGCATGCTGAAGCTGTACGCGCACAAGCTAGAATTTATGC 767

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     768 ACATTCTCACTCGCGTTAACAGGAAGGTGGCAACGGAATTCGAGTCCTTC 817

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSe 267
         |||:::|||:::|||||||||||||||||||||||||||||||||||||||
     818 TCCCTGGACTCCACTTTCCACGCAAAGAAACAGATCCCGTGTATTGTGTC 867

     267 rMetLeuThrLysGluLeuTyrPheTyrHis 277
         ||||||||||||||||||||||||||||||
     868 CATGCTCACGAAAGAACTGTACTTTTATCAC 898

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-1

seq_documentation_block:
; Sequence 1, Application US/08462969B
; Patent No. 6087150
;  GENERAL INFORMATION:
;    APPLICANT:  He, Wei-Wu et al.
;    TITLE OF INVENTION:  Interleukin-1 Beta Converting Enzyme
;    TITLE OF INVENTION:  Like Apoptosis Protease 3 and 4
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Ave.
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,969B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/334,251
;      FILING DATE:  11-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
```

```
;      REFERENCE/DOCKET NUMBER:  PF140P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1371 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-462-969B-1


alignment_scores:
            Quality:  712.00              Length:    284
              Ratio:  3.327                 Gaps:      4
 Percent Similarity:  75.352   Percent Identity: 51.056

alignment_block:
 US-09-613-508-4 x US-08-462-969B-1   ..

 Align seg 1/1  to: US-08-462-969B-1  from: 1  to: 1371

       5 GluAsnSerValAspSerLysSerIleLys...AsnLeuGluProLysIl 20
         |||:::|||||||||:::|||      :::   ::::::   |||:::::
     394 GAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCT 443

      20 eIleHisGlySerGluSerMetAspSerGlyIleSerLeuAspAsnSer. 36
         ::::      :::::::::      :::      |||:::   ::::::
     444 CTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCC 493

      37 .....TyrLysMetAspTyrProGluMetGly 45
                |||:::|||::::::   ::::::|||
     494 GGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGC 543

      46 LeuCysIleIleIleAsnAsnLysAsnPheHisLysSerThrGlyMetTh 62
         |||||||||||||||||||||||||||   |||   |||||||||
     544 AAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGG 593

      62 rSerArgSerGlyThrAspValAspAlaAlaAsnLeuArgGluThrPheA 79
         |||:::|||||||||   ||||||      |||   :::   ||||
     594 CGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCC 643

      79 rgAsnLeuLysTyrGluValArgAsnLysAsnAspLeuThrArgGluGlu 95
         ||:::|||   ::::::|||         ||||||   :::      :::
     644 GAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAG 693

      96 IleValGluLeuMetArgAspValSerLysGluAspHisSerLysArgSe 112
         :::   :::|||::::::   :::|||:::||||||||||::::::   ::
     694 ATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGC 743

     112 rSerPheValCysValLeuLeuSerHisGlyGluGluGlyIleIlePheG 129
         :   |||:::|||:::|||||||||||||||||||||:::::::|||:::|
     744 CTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATG 793

     129 lyThrAsnGlyProValAspLeuLysLysIleThrAsnPhePheArgGly 145
         ||   :::|||   :::   :::|||   :::|||      |||||||||
     794 GGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGG 843

     146 AspArgCysArgSerLeuThrGlyLysProLysLeuPheIleIleGlnAl 162
         |||||||||:::::::|||      |||||||||||||||:::|||||||
     844 GATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGC 893

     162 aCysArgGlyThrGluLeuAspCysGlyIleGluThrAspSerGlyValA 179
         ||||||||||||||||||||   ||||||::::::|||||||||
     894 TTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCA 943

     179 spAspAspMetAlaCys.........HisLysIleProValGluAlaAsp 192
           :::|||      :::         :::|||||||||||||||||||
     944 TCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGAC 993

     193 PheLeuTyrAlaTyrSerThrAlaProGlyTyrTyrSerTrpArgAsnSe 209
         ||||||:::|||||||||||||:::|||||||||||||||||||||||:::
     994 TTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCC 1043

     209 rLysAspGlySerTrpPheIleGlnSerLeuCysAlaMetLeuLysGlnT 226
         ||||||||||||:::|||:::||||||:::::|||:::::::
    1044 AGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGC 1093

     226 yrAlaAspLysLeuGluPheMetHisIleLeu.ThrArgValAsnArgLy 242
         :::::      ||||||:::|||:::|||||| ||||||||||   ::
    1094 ACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGACAG 1143

     242 sValAlaThrGluPheGluSerPheSerPheAspAlaThrPheHisAlaL 259
         :||||||   :::|||||||||   |||   |||      ||||||   |
    1144 AGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGA 1193

     259 ysLysGlnIleProCysIleValSerMetLeuThrLysGluLeuTyrPhe 275
         ||||||||||||||||||:::|||||||||||||||||||||||||||||
    1194 AGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTC 1243

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-446-925-4

seq_documentation_block:
; Sequence 4, Application US/08446925
; Patent No. 5672500
;  GENERAL INFORMATION:
;    APPLICANT:  Litwack, Gerald
;    APPLICANT:  Alnemri, Emad S.
;    APPLICANT:  Fernandez-Alnemri, Teresa
;    TITLE OF INVENTION:  Mch2, AN APOPTOTIC CYSTEINE
;    TITLE OF INVENTION:  PROTEASE,
;    TITLE OF INVENTION:  AND COMPOSITIONS FOR MAKING AND
;    TITLE OF INVENTION:  METHODS
;    TITLE OF INVENTION:  OF USING THE SAME
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewicz &
;      ADDRESSEE:  No. 5672500ris
;      STREET:  One Liberty Place, 46th floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/446,925
;      FILING DATE:
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeLuca, Mark
;      REGISTRATION NUMBER:  33,229
;      REFERENCE/DOCKET NUMBER:  TJU-1508
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1545 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  79..957
US-08-446-925-4


alignment_scores:
            Quality:  517.50              Length:    253
```

```
            Ratio:  2.891              Gaps:      1
 Percent Similarity:  70.751   Percent Identity: 41.107

alignment_block:
 US-09-613-508-4 x US-08-446-925-4   ..

 Align seg 1/1  to: US-08-446-925-4  from: 1  to: 1545

      37 TyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsnLy 53
         ||||||||||||:::   :::   |||:::::::::|||:::|||:::::
     187 TACAAAATGGACCACAGGAGGAGAGGAATTGCTTTAATCTTCAATCATGA 236

      53 sAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspValA 70
         ::::|||              :::   :::|||      |||   :::|
     237 GAGGTTCTTTTGGCACTTAACACTGCCAGAAAGGCGGCGCACCTGCGCAG 286

      70 spAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValArg 86
         ||      ||||||   :::   |||   :::|||   :::|||||||:::
     287 ATAGAGACAATCTTACCCGCAGGTTTTCAGATCTAGGATTTGAAGTGAAA 336

      87 AsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAspVa 103
            |||||||||      ||||||:::::      ::::::::::||
     337 TGCTTTAATGATCTTAAAGCAGAAGAACTACTGCTCAAAATTCATGAGGT 386

     103 lSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLeuS 120
         ||||      :::|||:::      :::   ||||||||||||:::||||
     387 GTCAACTGTTAGCCACGCAGATGCCGATTGCTTTGTGTGTGTCTTCCTGA 436

     120 erHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspLeu 136
         ||||||||||   :::   |||::::::   ::::::   :::::::::
     437 GCCATGGCGAAGGCAATCACATTTATGCATATGATGCTAAAATCGAAATT 486

     137 LysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThrGl 153
         :::   :::|||:::::|||:::|||||||:::|||:::||||||:::||
     487 CAGACATTAACTGGCTTGTTCAAAGGAGACAAGTGTCACAGCCTGGTTGG 536

     153 yLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAspC 170
         |||||||||:::|||||||||||||||||||||||||::::::   |||
     537 AAAACCCAAGATATTTATCATCCAGGCATGTCGGGGAAACCAGCACGATG 586

     170 ysGlyIle.....................................Glu 173
            :::                                      :::
     587 TGCCAGTCATTCCTTTGGATGTAGTAGATAATCAGACAGAGAAGTTGGAC 636

     174 ThrAspSerGlyValAspAspAspMetAlaCysHisLysIleProValGl 190
         |||:::          |||      :::   :::   :::|||:::
     637 ACCAACATAACTGAGGTGGATGCAGCCTCCGTTTACACGCTGCCTGCTGG 686

     190 uAlaAspPheLeuTyrAlaTyrSerThrAlaProGlyTyrTyrSerTrpA 207
          ||||||||||||   :::|||||:::|||   ||||||||||||   |
     687 AGCTGACTTCCTCATGTGTTACTCTGTTGCAGAAGGATATTATTCTCACC 736

     207 rgAsnSerLysAspGlySerTrpPheIleGlnSerLeuCysAlaMetLeu 223
         ||:::::::  :::|||||||||:::|||||:::|||||   ||||||
     737 GGGAAACTGTGAACGGCTCATGGTACATTCAAGATTTGTGTGAGATGTTG 786

     224 LysGlnTyrAlaAspLysLeuGluPheMetHisIleLeuThrArgValAs 240
            :::|||:::::::::|||||||||   ::::::|||||   |||||
     787 GGAAAATATGGCTCCTCCTTAGAGTTCACAGAACTCCTCACACTGGTGAA 836

     240 nArgLysValAlaThrGluPheGluSerPheSerPheAspAlaThrPheH 257
         ||||||||||:::   :::      :::|||      |||   :::
     837 CAGGAAAGTTTCTCAGCGCCGAGTGGACTTTTGCAAAGACCCAAGTGCAA 886

     257 isAlaLysLysGlnIleProCysIleValSerMetLeuThrLysGluLeu 273
           :::|||||||||:::|||||:::::::|||||||||||||||:::|||
     887 TTGGAAAGAAGCAGGTTCCCTGTTTTGCCTCAATGCTAACTAAAAAGCTG 936

     274 TyrPheTyr 276
         :::|||:::
     937 CATTTCTTT 945

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-146-331-4

seq_documentation_block:
; Sequence 4, Application US/09146331
; Patent No. 5958720
;  GENERAL INFORMATION:
;    APPLICANT:  Litwack, Gerald
;    APPLICANT:  Alnemri, Emad S.
;    APPLICANT:  Fernandez-Alnemri, Teresa
;    TITLE OF INVENTION:  Mch2, AN APOPTOTIC CYSTEINE
;    TITLE OF INVENTION:  PROTEASE,AND COMPOSITIONS FOR MAKING AND
;    TITLE OF INVENTION:  METHODS OF USING THE SAME
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewicz &
;      ADDRESSEE:  No. 5958720ris
;      STREET:  One Liberty Place, 46th floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/146,331
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/896,885
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeLuca, Mark
;      REGISTRATION NUMBER:  33,229
;      REFERENCE/DOCKET NUMBER:  TJU-1508
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1545 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  79..957
US-09-146-331-4


alignment_scores:
            Quality:  517.50              Length:    253
              Ratio:  2.891                 Gaps:      1
 Percent Similarity:  70.751   Percent Identity: 41.107

alignment_block:
 US-09-613-508-4 x US-09-146-331-4   ..

 Align seg 1/1  to: US-09-146-331-4  from: 1  to: 1545

      37 TyrLysMetAspTyrProGluMetGlyLeuCysIleIleIleAsnAsnLy 53
         ||||||||||||:::   :::   |||:::::::::|||:::|||:::::
     187 TACAAAATGGACCACAGGAGGAGAGGAATTGCTTTAATCTTCAATCATGA 236

      53 sAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAspValA 70
         ::::|||              :::   :::|||      |||   :::|
     237 GAGGTTCTTTTGGCACTTAACACTGCCAGAAAGGCGGCGCACCTGCGCAG 286

      70 spAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValArg 86
         ||      ||||||   :::   |||   :::|||   :::|||||||:::
     287 ATAGAGACAATCTTACCCGCAGGTTTTCAGATCTAGGATTTGAAGTGAAA 336
```

```
 87 AsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArgAspVa 103
337 TGCTTTAATGATCTTAAAGCAGAAGAACTACTGCTCAAAATTCATGAGGT 386

103 lSerLysGluAspHisSerLysArgSerSerPheValCysValLeuLeuS 120
387 GTCAACTGTTAGCCACGCAGATGCCGATTGCTTTGTGTGTGTCTTCCTGA 436

120 erHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValAspLeu 136
437 GCCATGGCGAAGGCAATCACATTTATGCATATGATGCTAAAATCGAAATT 486

137 LysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeuThrGl 153
487 CAGACATTAACTGGCTTGTTCAAAGGAGACAAGTGTCACAGCCTGGTTGG 536

153 yLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAspC 170
537 AAAACCCAAGATATTTATCATCCAGGCATGTCGGGGAAACCAGCACGATG 586

170 ysGlyIle.....................................Glu 173
587 TGCCAGTCATTCCTTTGGATGTAGTAGATAATCAGACAGAGAAGTTGGAC 636

174 ThrAspSerGlyValAspAspAspMetAlaCysHisLysIleProValGl 190
637 ACCAACATAACTGAGGTGGATGCAGCCTCCGTTTACACGCTGCCTGCTGG 686

190 uAlaAspPheLeuTyrAlaTyrSerThrAlaProGlyTyrTyrSerTrpA 207
687 AGCTGACTTCCTCATGTGTTACTCTGTTGCAGAAGGATATTATTCTCACC 736

207 rgAsnSerLysAspGlySerTrpPheIleGlnSerLeuCysAlaMetLeu 223
737 GGGAAACTGTGAACGGCTCATGGTACATTCAAGATTTGTGTGAGATGTTG 786

224 LysGlnTyrAlaAspLysLeuGluPheMetHisIleLeuThrArgValAs 240
787 GGAAAATATGGCTCCTCCTTAGAGTTCACAGAACTCCTCACACTGGTGAA 836

240 nArgLysValAlaThrGluPheGluSerPheSerPheAspAlaThrPheH 257
837 CAGGAAAGTTTCTCAGCGCCGAGTGGACTTTTGCAAAGACCCAAGTGCAA 886

257 isAlaLysLysGlnIleProCysIleValSerMetLeuThrLysGluLeu 273
887 TTGGAAAGAAGCAGGTTCCCTGTTTTGCCTCAATGCTAACTAAAAAGCTG 936

274 TyrPheTyr 276
937 CATTTCTTT 945
```

```
OM of: US-09-613-508-4 to: EST:*   out_format : pfs

Date: Aug 15, 2002  8:00 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO_spool/US09613508/runat_14082002_112647_7211/app_query.fasta_1.701
 -DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
 -MINLEN=0 -MAXLEN=2000000000 -USER=US09613508_@CGN1_1_3002
 -NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-613-508-4
Query length: 277
Database: EST:*
Database sequences: 13736207
Database length: -1841457050
Search time (sec): 2163.950000



score_list:
Sequence              Strd Orig    ZScore    EScore Len  1 Documentation  ..
gb_est2:BM461940       +  1399.00 3089.39  7.8e-163  913  ! BM461940 AGENCOURT_6418271 NIH_
gb_est1:AL534526       +  1347.00 2974.70  1.9e-156  827  ! AL534526 AL534526 LTI_FL013_FBr
gb_est1:AU119346       +  1236.00 2727.13  1.2e-142  846  ! AU119346 AU119346 HEMBA1 Homo s
gb_est2:BE780308       +  1152.00 2539.11  3.5e-132  911  ! BE780308 601468289F1 NIH_MGC_67
gb_est2:BI652204       +  1151.00 2538.50  3.8e-132  794  ! BI652204 603299029F1 NIH_CGAP_M
gb_est2:BI080567       +  1104.50 2433.72  2.6e-126  878  ! BI080567 602877052F1 NCI_CGAP_M
gb_est1:AU127028       +  1025.00 2258.44  1.5e-116  751  ! AU127028 AU127028 NT2RP2 Homo s
gb_est2:BI657919       +  1017.00 2242.26  1.2e-115  653  ! BI657919 603282147F1 NIH_CGAP_M
gb_est1:AU133051       +  1009.50 2221.12  1.8e-114  951  ! AU133051 AU133051 NT2RP4 Homo s
gb_est2:BI691090       +  1004.00 2213.57  4.7e-114  638  ! BI691090 603309866F1 NCI_CGAP_M
gb_est2:BG865023       +   930.00 2047.61  8.3e-105  700  ! BG865023 602799491F1 NIH_CGAP_M
gb_est2:BF966141       +   816.00 1795.40  9.3e-91   602  ! BF966141 602286412F1 NIH_MGC_95
gb_est1:AU125557       +   795.00 1744.99  6.0e-88   819  ! AU125557 AU125557 NT2RM4 Homo s
gb_est1:AW951022       +   794.50 1747.56  4.3e-88   599  ! AW951022 EST363092 MAGE reseque
gb_est2:BE284569       +   742.00 1628.67  1.8e-81   705  ! BE284569 601087184F1 NCI_CGAP_M
gb_est1:AU127381       +   742.00 1626.81  2.3e-81   826  ! AU127381 AU127381 NT2RP2 Homo s
gb_est1:AA011446       -   739.00 1627.40  2.1e-81   445  ! AA011446 zi01g12.s1 Soares_feta
gb_est1:AL542189       +   725.00 1587.00  3.8e-79   973  ! AL542189 AL542189 LTI_FL002_PL1
gb_est1:AA011445       +   704.00 1549.61  4.6e-77   438  ! AA011445 zi01g12.r1 Soares_feta
gb_est2:BI153295       +   702.00 1538.03  2.0e-76   802  ! BI153295 602917153F1 NCI_CGAP_I
gb_est2:BG170404       +   691.00 1511.35  6.2e-75   965  ! BG170404 602322734F1 NIH_MGC_89
gb_est2:T10341         +   652.00 1434.86  1.1e-70   399  ! T10341 seq1254 b4HB3MA Cot8-HAF
gb_est1:AV649560       +   651.00 1427.36  2.9e-70   624  ! AV649560 AV649560 GLC Homo sapi
gb_est2:BG287359       +   649.00 1417.27  1.1e-69   1007  ! BG287359 602382123F1 NIH_MGC_9
gb_est1:AV649613       +   648.00 1420.68  6.9e-70   624  ! AV649613 AV649613 GLC Homo sapi
gb_est2:BG177093       +   635.50 1387.93  4.6e-68   946  ! BG177093 602313329F1 NIH_MGC_85
gb_est1:AU127408       +   635.00 1394.83  1.9e-68   479  ! AU127408 AU127408 NT2RP2 Homo s
gb_est2:BI685778       +   598.00 1306.50  1.6e-63   790  ! BI685778 603309431F1 NCI_CGAP_M
gb_est2:BF971443       +   595.50 1303.12  2.4e-63   656  ! BF971443 602272969F1 NIH_MGC_84
gb_est2:H29199         +   589.00 1293.32  8.6e-63   441  ! H29199 ym31f11.r1 Soares infant
gb_est1:AL602756       +   588.00 1288.17  1.7e-62   565  ! AL602756 DKFZp686H1516_r1 686 (
gb_est2:BG025549       +   580.50 1266.93  2.5e-61   830  ! BG025549 602274928F1 NIH_MGC_85
gb_est1:AA984697       +   556.00 1218.39  1.3e-58   497  ! AA984697 am89h06.s1 Stratagene
gb_est1:AL602306       +   556.00 1217.79  1.4e-58   523  ! AL602306 DKFZp686J0612_r1 686 (
gb_est2:BG264309       +   548.00 1199.10  1.5e-57   563  ! BG264309 de76e01.y1 Kirschner e
gb_est2:BI821936       +   537.00 1168.73  7.5e-56   926  ! BI821936 603039745F1 NIH_MGC_11
gb_est2:BI132775       +   531.00 1164.26  1.3e-55   435  ! BI132775 AR035B10PBC05CB10S Por
gb_est2:BM459840       +   530.00 1151.38  6.9e-55   1075  ! BM459840 AGENCOURT_6422147 NIH
gb_est1:AW542198       -   517.00 1133.99  6.4e-54   402  ! AW542198 C0154G07-3 NIA Mouse E
gb_est2:BI692031       +   516.00 1121.23  3.3e-53   984  ! BI692031 603311680F1 NCI_CGAP_M
gb_htc:AK011710        +   516.00 1119.40  4.2e-53   1149  ! AK011710 Mus musculus 10 d
gb_est2:BG966729       +   512.00 1115.14  7.2e-53   774  ! BG966729 602834592F1 NCI_CG
gb_est2:BM149220       +   511.00 1120.28  3.7e-53   414  ! BM149220 TCAAP2Q11479 Pedia
gb_est2:BI824214       +   505.00 1100.00  5.0e-52   745  ! BI824214 603040584F1 NIH_MG
gb_est1:AI315897       +   504.50 1098.24  6.3e-52   787  ! AI315897 uj27d07.y1 Sugano

seq_name: gb_est2:BM461940

seq_documentation_block:
LOCUS       BM461940               913 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6418271 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5534002
            5', mRNA sequence.
ACCESSION   BM461940
VERSION     BM461940.1  GI:18510980
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 913)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12219  row: m  column: 11
            High quality sequence stop: 671.
FEATURES             Location/Qualifiers
     source          1. .913
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5534002"
                     /clone_lib="NIH_MGC_71"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
BASE COUNT      285 a    159 c    203 g    265 t      1 others
ORIGIN


alignment_scores:
            Quality: 1399.00              Length:    277
              Ratio:  5.125                 Gaps:      2
 Percent Similarity:  98.556   Percent Identity: 97.834

alignment_block:
 US-09-613-508-4 x BM461940   ..

 Align seg 1/1  to: BM461940  from: 1  to: 913

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      39 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 88

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      89 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 138

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     139 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 188

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
    189 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 238

     67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    239 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 288

     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    289 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 338

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    339 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 388

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    389 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 438

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    439 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 488

    151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    489 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 538

    167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    539 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 588

    184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    589 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 638

    201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
        |||||||||||||||||||||||:::|||||||||||||||||||||||||
    639 CCTGGTTATTATTCTTGGCGAAATTNCAAGGATGGCTCCTGGTTCATCCA 688

    217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    689 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 738

    234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    739 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTT 788

    251 SerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleVal.S 267
        |||||||||||||||||   ||||||:::|||||||||||||||||| |
    789 TCCTTTGACGCTACTTTTTCTGCCAAAGAACAGATTCCATGTATTGTTTT 838

    267 erMetLeu.ThrLysGluLeuTyrPhe 275
        ||||||||   ||||||||||||||||
    839 CCATGCTCCCCAAAGGAACTCTATTTT 865

seq_name: gb_est1:AL534526

seq_documentation_block:
LOCUS       AL534526                827 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL534526 LTI_FL013_FBrn1 Homo sapiens cDNA clone CS0DF004YA06 5
            prime, mRNA sequence.
ACCESSION   AL534526
VERSION     AL534526.1  GI:12798019
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 827)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .827
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DF004YA06"
                     /clone_lib="LTI_FL013_FBrn1"
                     /dev_stage="pooled tissue from post conception fetuses (20
                     week, 24 week and 26 week)"
                     /lab_host="DH10B"
                     /note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand
                     cDNA was primed with a NotI-oligo(dT) primer. Five prime
                     end enriched, double-stranded cDNA was digested with Not I
                     and cloned into the Not I and Eco RV sites of the
                     pCMVSPORT 6 vector.  Library was constructed by Life
                     Technologies.  Contact : Feng Liang Life Technologies, a
                     division of Invitrogen 9800 Medical Center Drive Rockville
                     , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
                     fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      266 a    140 c    192 g    229 t
ORIGIN


alignment_scores:
            Quality: 1347.00              Length:    259
              Ratio:  5.221                Gaps:      1
 Percent Similarity:  99.614   Percent Identity: 99.614

alignment_block:
 US-09-613-508-4 x AL534526   ..

 Align seg 1/1  to: AL534526  from: 1  to: 827

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
      51 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 100

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     101 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 150

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     151 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 200

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     201 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 250

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     251 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 300

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     301 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 350

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     351 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 400

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     401 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 450

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     451 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 500

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         |||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     501 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 550

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     551 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 600

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     601 GTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 650

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     651 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCA 700

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     701 GTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGC 750

     234 isIleLeuThrArgValAsnArgLysValAlaThrGlu.PheGluSerPh 250
         |||||||||||||||||||||||||||||||||||| ||||||||||||
     751 ACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTTGAGTCCTT 800

     250 eSerPheAspAlaThrPheHisAla 258
         |||||||||||||||||||||||||
     801 TTCCTTTGACGCTACTTTTCATGCA 825

seq_name: gb_est1:AU119346

seq_documentation_block:
LOCUS       AU119346                846 bp    mRNA    linear   EST 19-OCT-2000
DEFINITION  AU119346 HEMBA1 Homo sapiens cDNA clone HEMBA1005587 5', mRNA
            sequence.
ACCESSION   AU119346
VERSION     AU119346.1  GI:10934581
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 846)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .846
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEMBA1005587"
                     /clone_lib="HEMBA1"
                     /tissue_type="whole embryo, mainly head"
                     /dev_stage="embryo, 10 weeks"
                     /note="Vector: pME18SFL3"
BASE COUNT      266 a    150 c    200 g    226 t      4 others
ORIGIN


alignment_scores:
            Quality: 1236.00             Length:    267
              Ratio:  4.964                Gaps:      4
 Percent Similarity:  93.258   Percent Identity: 91.760
```

```
alignment_block:
 US-09-613-508-4 x AU119346   ..

 Align seg 1/1  to: AU119346  from: 1  to: 846

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      68 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 117

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     118 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 167

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     168 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 217

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     218 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 267

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     268 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 317

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||:::||||||||||||||||||||||||||||||||||
     318 AAGTCAGGAATAAAAGTGATCTTACACGTGAAGAAATTGTGGAATTGATG 367

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     368 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 417

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         * ||||||||||||||||||||||||||||||||||||||||||||||||||
     418 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 467

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     468 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 517

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     518 CTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGA 567

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     568 ACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGT 617

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         |||||||||||||||||||||||||||||||||   ||||||||||||||||
     618 GTCATAAAATACCAGTGGAGGCCGACTTCTTGNATGCATACTCCACAGCA 667

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         |||||||||||||||||||||||||||||||||   |||   |||||||||
     668 CCTGGNTATTATTCTTGGCGAAATTCAAAGGATGGCTTCTGGNTCATCCA 717

     217 nSerLeuCysAlaMetLeuLysGlnTyrAla.AspLysLeuGluPheMet 233
         |||||||||||||||||||||||||||||| ||||||||||||||||||||
     718 GTCGCTTTGTGCCATGCTGAAACAGTATGCCCGACAAGCTTGAATTTATG 767

     234 HisIleLeuThrArgVal.....,.....AsnArgLysValAlaThrGl 246
         ||||||||||||||||||           ||||||      ::::::::
     768 CACATTCTTACCCGGGTTACCGAANGTGGCAACAGAACTTGGTCCTCTTC 817

     246 uPheGluSerPheSerPheAspAlaThr.PheHisAlaLysLysGln 261
         :|||                   |||    ||||||||||||||||
     818 CTTT.................GCGCTACTTTCATGCCAAGAAACAG 846

seq_name: gb_est2:BE780308

seq_documentation_block:
```

```
LOCUS       BE780308                911 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601468289F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3871390 5',
            mRNA sequence.
ACCESSION   BE780308
VERSION     BE780308.1  GI:10201506
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 911)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9624  row: a  column: 23
            High quality sequence stop: 768.
FEATURES             Location/Qualifiers
     source          1. .911
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3871390"
                     /clone_lib="NIH_MGC_67"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      287 a    175 c    205 g    244 t
ORIGIN


alignment_scores:
           Quality: 1152.00            Length:    280
             Ratio:  4.553              Gaps:      7
 Percent Similarity:  90.357   Percent Identity: 86.786

alignment_block:
 US-09-613-508-4 x BE780308    ..

 Align seg 1/1  to: BE780308  from: 1  to: 911

       1 MetGluAsnThr.GluAsnSerValAspSerLysSerIleLysAsnLeuG 17
         |||||||||||| ||||||||||||||||||||||||||||||||||||
       1 ATGGAGAACACTGGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGG 50

      17 luProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeu 33
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 AACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTG 100

      34 AspAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIl 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 GACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAAT 150

      50 eAsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyT 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTA 200

      67 hrAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyr 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 CAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATAT 250

      84 GluValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMe 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||

     251 GAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGAT 300

     100 tArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysV 117
         |||||||||||||||||||||||||||||||||||   ||||||||||
     301 GCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAG.TTTGTTTGTG 349

     117 alLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyPro 133
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     350 TGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCT 399

     134 ValAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSe 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     400 GTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAG 449

     150 rLeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrG 167
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     450 TCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAG 499

     167 luLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAla 183
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     500 AACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCG 549

     184 CysHisLysIleProValGluAlaAspPhe.LeuTyrAlaTyrSerThrA 200
         |||||||||||||||||||||||||||| ||||||||||||||||||||
     550 TGTCATAAAATACCAGTGGAGGCCGACTTCTTTGTATGCATACTCCACAG 599

     200 laProGly...TyrTyrSerTrp.ArgAsnSerLysAspGlySerTrpPh 215
         ||||||||   :::      ||| |||:::          ||||||||
     600 CACCTGGGTTATTTATTCTATGGCCGAAAATTCACAGGATGGTCCTGGTT 649

     215 eIleGlnSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGlu. 231
         ||||||||||   ||||||||||||||||||||||||||||||||||||
     650 CATCCAGTCGGCTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAAA 699

     232 ..PheMetHisIleLeuThrArgValAsnArgLysVal.AlaThr.GluP 247
           :::      :::||||||||||||||||||:::||| |||||| ||||
     700 TTTATGCAACATTTCTTACCCGGGTTAACCGACAGGTGGGCAACCAGAAT 749

     247 heGluSerPheSerPheAspAlaThrPheHisAlaLysLysGlnIlePro 263
         ||||||||||||||:::          |||||||||::::::
     750 TTGAGTCCTTTTCCTTGACGCTACTTTTCCATGCAGAAGAAACAGATGCA 799

     264 CysIleValSerMetLeuThrLysGluLeu 273
         |||:::          :::   |||   |||
     800 TGTTTGTTTCCATGCTTCCCAAAGACTCTC 829

seq_name: gb_est2:BI652204

seq_documentation_block:
LOCUS       BI652204                794 bp    mRNA    linear   EST 12-SEP-2001
DEFINITION  603299029F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5339623 5',
            mRNA sequence.
ACCESSION   BI652204
VERSION     BI652204.1  GI:15566440
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 794)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11862  row: j  column: 08
```

```
          High quality sequence start: 37
          High quality sequence stop: 788.
FEATURES             Location/Qualifiers
     source          1. .794
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5339623"
                     /clone_lib="NIH_CGAP_Mam3"
                     /tissue_type="tumor, gross tissue"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert 2 kb. Library constructed by Life
                     Technologies, catalog #12017-018. Investigators providing
                     samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference
                     for transgenic model: Xu et al., Nature Genetics 22, 37-43
                     (1999). Note: this is a NCI_CGAP Library."
BASE COUNT      225 a    164 c    209 g    196 t
ORIGIN


alignment_scores:
            Quality: 1151.00             Length:    257
              Ratio:  4.679               Gaps:      2
 Percent Similarity:  95.720   Percent Identity: 84.047

alignment_block:
 US-09-613-508-4 x BI652204   ..

 Align seg 1/1  to: BI652204  from: 1  to: 794

      23 GlySerGluSerMetAspSerGlyIle.SerLeuAspAsnSerTyrLysM 39
         |||::::::::::|||||:::|||:::    ||||||:::|||||||:::|
      17 GGAGCAAGTCAGATGGACGCTGGGATGCTATCTGGACAGTAGTTACAGAA 66

      39 etAspTyrProGluMetGlyLeuCysIleIleIleAsnAsnLysAsnPhe 55
         ||||||||||||||||||||:::|||||||||   ||||||||||||||||
      67 TGGATTATCCTGAAATGGGCATATGCATAATAAAAAATAATAAGAACTTC 116

      56 HisLysSerThrGlyMetThrSerArgSerGlyThrAspValAspAlaAl 72
         ||||||||||||||||||:::||||||||||||||||||||||||||||||
     117 CATAAGAGCACTGGAATGTCATCTCGCTCTGGTACGGATGTGGACGCAGC 166

      72 aAsnLeuArgGluThrPheArgAsnLeuLysTyrGluValArgAsnLysA 89
         ||||||||||||||||||   :::||||||||||:::|||||||||||||
     167 CAACCTCAGAGAGACATTCATGGGCCTGAAATACCAAGTCAGGAATAAAA 216

      89 snAspLeuThrArgGluGluIleValGluLeuMetArgAspValSerLys 105
         ||||||||||||||||||:::|||:::||||||||||   :::||||||||
     217 ATGATCTTACTCGTGAAGACATTTTGGAATTAATGGATAGTGTTTCTAAG 266

     106 GluAspHisSerLysArgSerSerPheValCysValLeuLeuSerHisGl 122
         ||||||||||||||||||||||||||||||||||:::||||||||||||
     267 GAAGATCATAGCAAAAGGAGCAGCTTTGTGTGTGTGATTCTAAGCCATGG 316

     122 yGluGluGlyIleIlePheGlyThrAsnGlyProValAspLeuLysLysI 139
         |:::||||||:::|||:::|||||||||||||||||||:::||||||||||:
     317 TGATGAAGGGGTCATTTATGGGACAAATGGGCCTGTTGAACTGAAAAAGT 366

     139 leThrAsnPhePheArgGlyAspArgCysArgSerLeuThrGlyLysPro 155
         ::|||:::|||||||||||||||   ||||||||||||||||||||||||
     367 TGACTAGCTTCTTCAGAGGCGACTACTGCCGGAGTCTGACTGGAAAGCCG 416

     156 LysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAspCysGlyIl 172
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     417 AAACTCTTCATCATTCAGGCCTGCCGGGGTACGGAGCTGGACTGTGGCAT 466

     172 eGluThrAspSerGlyValAspAspAspMetAlaCysHisLysIleProV 189
         ||||||||||||||||:::|||:::::::||||||||:::||||||||||
     467 TGAGACAGACAGTGGGACTGATGAGGAGATGGCTTGCCAGAAGATACCGG 516

     189 alGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGlyTyrTyrSer 205
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     517 TGGAGGCTGACTTCCTGTATGCTTACTCTACAGCACCTGGTTACTATTCC 566

     206 TrpArgAsnSerLysAspGlySerTrpPheIleGlnSerLeuCysAlaMe 222
         ||||||||||||||||||||||||||||||||||||||||||:::||
     567 TGGAGAAATTCAAAGGACGGGTCGTGGTTCATCCAGTCCCTTTGCAGCAT 616

     222 tLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleLeuThrArgV 239
         |||||||   ||||||   ||||||||||||||||||||||||||||||
     617 GCTGAAGCTGTACGCGCACAAGCTAGAATTTATGCACATTCTCACTCGCG 666

     239 alAsnArgLysValAlaThrGluPheGluSerPheSerPheAspAlaThr 255
         ||:::   :::   ||||||||||||||||||||||:::|||:::|||
     667 TAACAGGAAGGTGGGCAACGGAATTCGAGTCCTTCTCCCTGGACTCCACT 716

     256 PheHisAlaLysLysGlnIleProCysIleValSerMetLeu.ThrLysG 272
         |||||||||||||||||||||||||||||||||||||| |||||||
     717 TTCCACGCAAAGAAACAGATCCCGTGTATTGTGTCCATGCTTCACGAAAG 766

     272 luLeuTyrPheTyrHis 277
         ||||||||||||||||
     767 AACTGTACTTTTATCAC 783

seq_name: gb_est2:BI080567

seq_documentation_block:
LOCUS       BI080567                878 bp    mRNA    linear   EST 20-JUN-2001
DEFINITION  602877052F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5008572 5',
            mRNA sequence.
ACCESSION   BI080567
VERSION     BI080567.1  GI:14498897
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 878)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11053  row: h  column: 13
            High quality sequence stop: 794.
FEATURES             Location/Qualifiers
     source          1. .878
                     /organism="Mus musculus"
                     /strain="FVB/N-3"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5008572"
                     /clone_lib="NCI_CGAP_Mam2"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      266 a    174 c    227 g    211 t
ORIGIN


alignment_scores:
            Quality: 1104.50             Length:    278
              Ratio:  4.348               Gaps:      4
 Percent Similarity:  91.367   Percent Identity: 79.137
```

```
alignment_block:
 US-09-613-508-4 x BI080567   ..

 Align seg 1/1  to: BI080567  from: 1  to: 878

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         |||||||||:::::::::|||||||||||||||||||||||:::|||:::||
      41 ATGGAGAACAACAAAACCTCAGTGGATTCAAAATCCATTAATAATTTTGA 90

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         |   |||   ||||||||||||:::|||:::||||||||||||   ||||
      91 AGTAAAGACCATACATGGGAGCAAGTCAGTGGACTCTGGGATCTATCTGG 140

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||:::|||||||||||||||||||||||||||||:::||||||||||||
     141 ACAGTAGTTACAAAATGGATTATCCTGAAATGGGCATATGCATAATAATT 190

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||:::|||||||||||||||
     191 AATAATAAGAACTTCCATAAGAGCACTGGAATGTCATCTCGCTCTGGTAC 240

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         |||||||||||||||||||||||||||||||   :::||||||||||:
     241 GGATGTGGACGCAGCCAACCTCAGAGAGACATTCATGGGCCTGAAATACC 290

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ::|||||||||||||||||||||||||||||:::|||:::|||||||||
     291 AAGTCAGGAATAAAAATGATCTTACTCGTGAAGACATTTTGGAATTAATG 340

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
           :::|||||||||||||||||||||||||||||||||||||||||||||
     341 GATAGTGTTTCTAAGGAAGATCATAGCAAAAGGAGCAGCTTTGTGTGTGT 390

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         |:::|||||||||||:::|||||||:::|||:::|||||||||||||||
     391 GATTCTAAGCCATGGTGATGAAGGGGTCATTTATGGGACAAATGGGCCTG 440

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||:::|||||||||:::|||:::|||||||||||||||   |||||||||
     441 TTGAACTGAAAAAGTTGACTAGCTTCTTCAGAGGCGACTACTGCCGGAGT 490

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     491 CTGACTGGAAAGCCGAAACTCTTCATCATTCAGGCCTGCCGGGGTACGGA 540

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||:::|||:::::::|||||||
     541 GCTGGACTGTGGCATTGAGACAGACAGTGGGACTGATGAGGAGATGGCTT 590

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||:::||||||||||||||||||||||||||||||||||||||||||||
     591 GCCAGAAGATACCGGTGGAGGCTGACTTCCTGTATGCTTACTCTACAGCA 640

     201 ProGlyTyr.TyrSerTrpArg.AsnSerLysAspGlySerTrpPheIle 216
         |||||||||    |||||||| ||||||||||||||||||   ||||||
     641 CCTGGTTACTAATTCCTGGAGAAAATTCAAAGGACGGGTCGTGCTTCATC 690

     217 GlnSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMe 233
         ||||||||||||:::|||||||||   ||||||   |||||||||:::||
     691 CAGTCCCTTTGCAGCATGCTGAAGCTGTACGCGCACAAGCTAGAATTAAT 740

     233 tHisIleLeu.ThrArgValAsnArgLysValAlaThrGluPheGluSer 249
         |||||||||| ||||||||||||||||||   |||   ||||||||||||
     741 GCACATTCTCAACTCGCGTTAACAGGAAGGGGGCACCGGAATTCGAGTCC 790

     250 PheSerPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleVa 266
         ||||||:::|||::::::   :::||||||||||:::
     791 TTCTCCCTGGACTCCAATGTTCAAGCAAAGAAAAGA...TCCGTGTATGG 837

     266 lSerMetLeuThrLysGluLeuTyrPhe 275
          ||||||||     :::|||   |||
     838 GTCCATGCTCAGAACAAAACTGTCCTTT 865
```

```
seq_name: gb_est1:AU127028

seq_documentation_block:
LOCUS       AU127028                751 bp    mRNA    linear   EST 23-OCT-2000
DEFINITION  AU127028 NT2RP2 Homo sapiens cDNA clone NT2RP2000417 5', mRNA
            sequence.
ACCESSION   AU127028
VERSION     AU127028.1  GI:10951744
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 751)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .751
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP2000417"
                     /clone_lib="NT2RP2"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                     cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT      242 a    129 c    175 g    194 t     11 others
ORIGIN


alignment_scores:
           Quality: 1025.00             Length:    218
             Ratio:  4.928                Gaps:      2
 Percent Similarity:  95.413   Percent Identity: 93.578

alignment_block:
 US-09-613-508-4 x AU127028   ..

 Align seg 1/1  to: AU127028  from: 1  to: 751

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
      97 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 146

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     147 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 196

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     197 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 246

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     247 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 296

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     297 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 346
```

```
     84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    347 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 396

    101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    397 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 446

    117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
        |||||||:::||||||||||||||||||||||||||||||||||||||||
    447 GCTTCTGANCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 496

    134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
        |||||||||||||||||||||||||||||||   ||||||||   :::
    497 TTGACCTGAAAAAAATAACAAACTTTTTCAGANGGGATCGTTGTANAANT 546

    151 LeuThr.GlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrG 167
        |||||| ||||||||||||||||||||||   ||||||||||||||||
    547 CTAACTGGGAAAACCCAAACTTTTCATTATTCANGCCTGCCGTGGTACAG 596

    167 luLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAla 183
        ||||||||||||||||   ||||||||||||||||||||||||||||||
    597 AACTGGACTGTGGCATTGANACAGACAGTGGTGTTGATGATGACATGGCG 646

    184 CysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAl 200
        ||||||||||||||   |||:::|||||||||   ||||||||||||::
    647 TGTCATAAAATACCATG.GAAGNCGACTTCTTGTNTGCATACTCCACANC 695

    200 aProGlyTyrTyrSerTrp..ArgAsnSerLysAspGlySerTrpPheIl 216
        :|||||||||||||||     |||   ||||||||||||||||||||
    696 ACCTGGTTATTATTCTTGGGCNGAAATCCAAAGGATGGCTCCTGGTTCAT 745

    216 e 216
        |
    746 C 746

seq_name: gb_est2:BI657919

seq_documentation_block:
LOCUS       BI657919                653 bp    mRNA    linear   EST 12-SEP-2001
DEFINITION  603282147F1 NIH_CGAP_Mam4 Mus musculus cDNA clone IMAGE:5326643 5',
            mRNA sequence.
ACCESSION   BI657919
VERSION     BI657919.1  GI:15572155
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 653)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
            Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11828  row: m  column: 12
            High quality sequence stop: 650.
FEATURES             Location/Qualifiers
     source          1. .653
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5326643"
                     /clone_lib="NIH_CGAP_Mam4"
                     /tissue_type="tumor, gross tissue"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert 2.5 kb.  Library constructed by Life
                     Technologies, catalog # 12018-016. Investigators providing
                     samples: Lothar Hennighausen/Priscilla Furth, NIH
                     Reference for transgenic model: Li et al., Cell Growth and
                     Differentiation 7, 3-11 (1996). Note: this is a NCI_CGAP
                     Library."
BASE COUNT      178 a    140 c    173 g    162 t
ORIGIN


alignment_scores:
           Quality: 1017.00            Length:    218
             Ratio:  4.797               Gaps:      1
 Percent Similarity:  97.248   Percent Identity: 87.156

alignment_block:
 US-09-613-508-4 x BI657919   ..

 Align seg 1/1  to: BI657919  from: 1  to: 653

     52 AsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyThrAs 68
        ||||||||||||||||||||||||||||:::||||||:::||||||||
      2 AATAAGAACTTCCATAAGAGCACTGGAATGTCATCTCGCGCTGGTACGGA 51

     68 pValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrGluV 85
        |||||||||||||||||||||||||||||   :::|||||||||:::|
     52 TGTGGACGCAGCCAACCTCAGAGAGACATTCATGGGCCTGAAATACCAAG 101

     85 alArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMetArg 101
        |||||||||||||||||||||||||||||:::|||:::|||||||||
    102 TCAGGAATAAAAATGATCTTACTCGTGAAGACATTTTGGAATTAATGGAT 151

    102 AspValSerLysGluAspHisSerLysArgSerSerPheValCysValLe 118
        :::|||||||||||||||||||||||||||||||||||||||||||||::
    152 AGTGTTTCTAAGGAAGATCATAGCAAAAGGAGCAGCTTTGTGTGTGTGAT 201

    118 uLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProValA 135
        :|||||||||||:::||||||:::|||:::|||||||||||||||||||:
    202 TCTAAGCCATGGTGATGAAGGGGTCATTTATGGGACAAATGGGCCTGTTG 251

    135 spLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSerLeu 151
        ::|||||||||:::|||:::|||||||||||||||   ||||||||||||
    252 AACTGAAAAAGTTGACTAGCTTCTTCAGAGGCGACTACTGCCGGAGTCTG 301

    152 ThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGluLe 168
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    302 ACTGGAAAGCCGAAACTCTTCATCATTCAGGCCTGCCGGGGTACGGAGCT 351

    168 uAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaCysH 185
        |||||||||||||||||||||||||:::|||:::::|||||||||:
    352 GGACTGTGGCATTGAGACAGACAGTGGGACTGATGAGGAGATGGCTTGCC 401

    185 isLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAlaPro 201
        ::||||||||||||||||||||||||||||||||||||||||||||||||
    402 AGAAGATACCGGTGGAGGCTGACTTCCTGTATGCTTACTCTACAGCACCT 451

    202 GlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGlnSe 218
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    452 GGTTACTATTCCTGGAGAAATTCAAAGGACGGGTCGTGGTTCATCCAGTC 501

    218 rLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisI 235
        |||||||:::||||||||   ||||||   ||||||||||||||||||
    502 CCTTTGCAGCATGCTGAAGCTGTACGCGCACAAGCTAGAATTTATGCACA 551

    235 leLeuThrArgValAsnArgLys.ValAlaThrGluPheGluSerPheSe 251
        ||||||||||||||||||:: ||||||||||||||||||||||||||||
    552 TTCTCACTCGCGTTAACAGGAACGGTGGCAACGGAATTCGAGTCCTTCTC 601

    251 rPheAspAlaThrPheHisAlaLysLysGlnIleProCysIleValSerM 268
        |:::|||:::||||||||||||||||||||||||||||||||||||||||
```

```
     602 CCTGGACTCCACTTTCCACGCAAAGAAACAGATCCCGTGTATTGTGTCCA 651

     268 et 268
         ||
     652 TG 653

seq_name: gb_est1:AU133051

seq_documentation_block:
LOCUS       AU133051                951 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  AU133051 NT2RP4 Homo sapiens cDNA clone NT2RP4001162 5', mRNA
            sequence.
ACCESSION   AU133051
VERSION     AU133051.1  GI:10993590
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 951)
  AUTHORS   Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J.,
            Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuho,Y. and
            Isogai,T.
  TITLE     HRI human cDNA project (Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y.,
            Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T., Sugano
            ,S., Masuho,Y., Isogai,T.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .951
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP4001162"
                     /clone_lib="NT2RP4"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                     cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT      276 a    186 c    220 g    266 t      3 others
ORIGIN


alignment_scores:
            Quality: 1009.50             Length:    285
              Ratio:  4.467               Gaps:      7
 Percent Similarity:  79.298   Percent Identity: 76.491

alignment_block:
 US-09-613-508-4 x AU133051   ..

 Align seg 1/1  to: AU133051  from: 1  to: 951

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     236 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 285

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     286 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 335

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     336 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 385

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||::
     386 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTGC 435

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         :||||||||||||||||||||||||||||||||||||||||||||||||||
     436 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 485

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     486 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 535

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     536 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 585

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     586 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 635

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     636 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 685

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||
     686 CTAACTGGAAAACCCAAACTTTTCATTATTCA..................... 717

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184

     717 ..................................................... 717

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200

     717 ..................................................... 717

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGly.SerTrpPheIleG 217
         |||||||||||||||||||||||||||||||||| ||||||||||||||
     718 ...GGGTATTATTCTTGGCGAAATTCAAAGGATGGCTTCCTGGTTCATCC 764

     217 lnSerLeuCysAlaMetLeuLysGlnTyrAla..AspLysLeuGluPheM 233
         ||||||||||||||||||||||||||||::::  ||||||||||||||||
     765 AGTCGCTTTGTGCCATGCTGAAACAGTATTGCCCGACAAGCTTGAATTTA 814

     233 etHisIle.LeuThrArgValAsnArgLysValAlaThr.GluPheGluS 249
         ||   ||| |||||||||||||||||||||||||||| :::|||||||||
     815 TGCCCATTCTTAACCCGGGTTAACCGAAAGGTGGCAACNNAAATTTGAGT 864

     249 erPheSerPheAspAlaThrPheHisAlaLysLysGln............. 261
         ||:::   :::      |||||||||||||||:::
     865 CCCTTTTCCTTTGACCTACTTTTCATGCAAAGGAAACCGAATTCATGNAT 914

     262 ...IleProCysIleValSerMetLeuThrLysGluLeuTyrPheTyr 276
         :::||||||            :::   ||||||||||||||||||
     915 TGGGTTCCATGC............TTCCCAAAAGAACTCTATTTTTAT 950

seq_name: gb_est2:BI691090

seq_documentation_block:
LOCUS       BI691090                638 bp    mRNA    linear   EST 18-SEP-2001
DEFINITION  603309866F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5345633 5',
            mRNA sequence.
ACCESSION   BI691090
VERSION     BI691090.1  GI:15653719
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 638)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11878  row: d  column: 18
            High quality sequence stop: 636.
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5345633"
                     /clone_lib="NCI_CGAP_Mam6"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      178 a    135 c    168 g    157 t
ORIGIN


alignment_scores:
            Quality: 1004.00            Length:    212
              Ratio:  4.874              Gaps:      0
 Percent Similarity:  97.170   Percent Identity: 87.736

alignment_block:
 US-09-613-508-4 x BI691090    ..

 Align seg 1/1  to: BI691090  from: 1  to: 638

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||:::|||   ||||||||
       2 AATAATAAGAACTTCCATAAGAGCACTGGAATGTCATCTGGCTCTGGTAC 51

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||   :::||||||||||:
      52 GGATGTGGACGCAGCCAACCTCAGAGAGACATTCATGGGCCTGAAATACC 101

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ::|||||||||||||||||||||||||||||:::|||:::|||||||||
     102 AAGTCAGGAATAAAAATGATCTTACTCGTGAAGACATTTTGGAATTAATG 151

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         :::|||||||||||||||||||||||||||||||||||||||||||||||
     152 GATAGTGTTTCTAAGGAAGATCATAGCAAAAGGAGCAGCTTTGTGTGTGT 201

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         |:::|||||||||||||:::|||||||:::|||:::|||||||||||||||
     202 GATTCTAAGCCATGGTGATGAAGGGGTCATTTATGGGACAAATGGGCCTG 251

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||:::|||||||||:::|||:::|||||||||||||||   |||||||||
     252 TTGAACTGAAAAAGTTGACTAGCTTCTTCAGAGGCGACTACTGCCGGAGT 301

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     302 CTGACTGGAAAGCCGAAACTCTTCATCATTCAGGCCTGCCGGGGTACGGA 351

     167 uLeuAspCysGlyIleGluThrAspSerGlyValAspAspAspMetAlaC 184
         ||||||||||||||||||||||||||||||:::|||:::::|||||||
     352 GCTGGACTGTGGCATTGAGACAGACAGTGGGACTGATGAGGAGATGGCTT 401

     184 ysHisLysIleProValGluAlaAspPheLeuTyrAlaTyrSerThrAla 200
         ||:::|||||||||||||||||||||||||||||||||||||||||||||
     402 GCCAGAAGATACCGGTGGAGGCTGACTTCCTGTATGCTTACTCTACAGCA 451

     201 ProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerTrpPheIleGl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     452 CCTGGTTACTATTCCTGGAGAAATTCAAAGGACGGGTCGTGGTTCATCCA 501

     217 nSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeuGluPheMetH 234
         ||||||||||:::|||||||||    ||||||    ||||||||||||||||
     502 GTCCCTTTGCAGCATGCTGAAGCTGTACGCGCACAAGCTAGAATTTATGC 551

     234 isIleLeuThrArgValAsnArgLysValAlaThrGluPheGluSerPhe 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     552 ACATTCTCACTCGCGTTAACAGGAAGGTGGCAACGGAATTCGAGTCCTTC 601

     251 SerPheAspAlaThrPheHisAlaLysLysGlnIle 262
         |||:::|||:::|||||||||||||||||||||||||
     602 TCCCTGGACTCCACTTTCCACGCAAAGAAACAGATC 637

seq_name: gb_est2:BG865023

seq_documentation_block:
LOCUS       BG865023                 700 bp    mRNA    linear   EST 29-MAY-2001
DEFINITION  602799491F1 NIH_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4934806 5',
            mRNA sequence.
ACCESSION   BG865023
VERSION     BG865023.1  GI:14215561
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 700)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
            Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10864  row: f  column: 23
            High quality sequence stop: 700.
FEATURES             Location/Qualifiers
     source          1. .700
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4934806"
                     /clone_lib="NIH_CGAP_Mam4"
                     /tissue_type="tumor, gross tissue"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert 2.5 kb.  Library constructed by Life
                     Technologies, catalog # 12018-016. Investigators providing
                     samples: Lothar Hennighausen/Priscilla Furth, NIH
                     Reference for transgenic model: Li et al., Cell Growth and
                     Differentiation 7, 3-11 (1996). Note: this is a NCI_CGAP
                     Library."
BASE COUNT      184 a    147 c    190 g    179 t
ORIGIN


alignment_scores:
            Quality:  930.00            Length:    197
              Ratio:  4.844              Gaps:      0
 Percent Similarity:  97.462   Percent Identity: 88.832
```

```
alignment_block:
 US-09-613-508-4 x BG865023   ..

 Align seg 1/1  to: BG865023  from: 1  to: 700

      81 LeuLysTyrGluValArgAsnLysAsnAspLeuThrArgGluGluIleVa 97
         |||||||||:::|||||||||||||||||||||||||||||||||:::|||
       6 CTGAAATACCAAGTCAGGAATAAAAATGATCTTACTCGTGAAGACATTTG 54

      97 lGluLeuMetArgAspValSerLysGluAspHisSerLysArgSerSerP 114
         |||||||||   :::|||||||||||||||||||||||||||||||||||
      55 .GAATTAATGGATAGTGTTTCTAAGGAAGATCATAGCAAAAGGAGCAGCT 104

     114 heValCysValLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThr 130
         ||||||||||:::|||||||||||:::|||||:::|||:::|||||||
     105 TTGTGTGTGTGATTCTAAGCCATGGTGATGAAGGGGTCATTTATGGGACA 154

     131 AsnGlyProValAspLeuLysLysIleThrAsnPhePheArgGlyAspAr 147
         ||||||||||||:::|||||||||:::|||:::|||||||||||||||
     155 AATGGGCCTGTTGAACTGAAAAAGTTGACTAGCTTCTTCAGAGGCGACTA 204

     147 gCysArgSerLeuThrGlyLysProLysLeuPheIleIleGlnAlaCysA 164
         |||||||||||||||||||||||||||||||||||||||||||||||||
     205 CTGCCGGAGTCTGACTGGAAAGCCGAAACTCTTCATCATTCAGGCCTGCC 254

     164 rgGlyThrGluLeuAspCysGlyIleGluThrAspSerGlyValAspAsp 180
         |||||||||||||||||||||||||||||||||||||||||:::|||:::
     255 GGGGTACGGAGCTGGACTGTGGCATTGAGACAGACAGTGGGACTGATGAG 304

     181 AspMetAlaCysHisLysIleProValGluAlaAspPheLeuTyrAlaTy 197
         :::|||||||||:::|||||||||||||||||||||||||||||||||||
     305 GAGATGGCTTGCCAGAAGATACCGGTGGAGGCTGACTTCCTGTATGCTTA 354

     197 rSerThrAlaProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerT 214
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     355 CTCTACAGCACCTGGTTACTATTCCTGGAGAAATTCAAAGGACGGGTCGT 404

     214 rpPheIleGlnSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeu 230
         ||||||||||||||||||:::|||||||||   ||||||   ||||||
     405 GGTTCATCCAGTCCCTTTGCAGCATGCTGAAGCTGTACGCGCACAAGCTA 454

     231 GluPheMetHisIleLeuThrArgValAsnArgLysValAlaThrGluPh 247
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     455 GAATTTATGCACATTCTCACTCGCGTTAACAGGAAGGTGGCAACGGAATT 504

     247 eGluSerPheSerPheAspAlaThrPheHisAlaLysLysGlnIleProC 264
         ||||||||||||||:::|||:::|||||||||||||||||||||||||||
     505 CGAGTCCTTCTCCCTGGACTCCACTTTCCACGCAAAGAAACAGATCCCGT 554

     264 ysIleValSerMetLeuThrLysGluLeuTyrPheTyrHis 277
         |||||||||||||||||||||||||||||||||||||||
     555 GTATTGTGTCCATGCTCACGAAAGAACTGTACTTTTATCAC 595

seq_name: gb_est2:BF966141

seq_documentation_block:
LOCUS       BF966141                602 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  602286412F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4375323 5',
            mRNA sequence.
ACCESSION   BF966141
VERSION     BF966141.1  GI:12333356
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 602)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10040  row: g  column: 04
            High quality sequence stop: 599.
FEATURES             Location/Qualifiers
     source          1. .602
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4375323"
                     /clone_lib="NIH_MGC_95"
                     /tissue_type="hippocampus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.5 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for  full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health).  Note: this is a NIH_MGC Library."
BASE COUNT      206 a     97 c    149 g    150 t
ORIGIN


alignment_scores:
            Quality:  816.00             Length:    167
              Ratio:  4.945                Gaps:      0
 Percent Similarity:  98.802   Percent Identity: 97.605

alignment_block:
 US-09-613-508-4 x BF966141   ..

 Align seg 1/1  to: BF966141  from: 1  to: 602

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 150

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 200

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 250

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 300

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 350

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 400

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||::|||||||
     401 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTAGTTTGTGT 450

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         |||||||||||||||||||||||||   ||||||||||||||||||||
     451 GCTTCTGAGCCATGGTGAAGAAGGAATAAT.TTTGGAACAAATGGACCTG 499
```

```
     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     500 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGT 549

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         |||   ||||||   ||||||||||||||||||:::||||||||||||||
     550 CTACTGGGAAAAACCAAACTTTTCATTATTCAGGGCTGCCGTGGTACAGA 599

     167 u 167
         |
     600 A 600

seq_name: gb_est1:AU125557

seq_documentation_block:
LOCUS       AU125557                819 bp    mRNA    linear   EST 23-OCT-2000
DEFINITION  AU125557 NT2RM4 Homo sapiens cDNA clone NT2RM4001791 5', mRNA
            sequence.
ACCESSION   AU125557
VERSION     AU125557.1  GI:10950273
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 819)
  AUTHORS   Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,
            Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,
            Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki
            ,Y., Sugano,S., Isogai,T.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .819
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RM4001791"
                     /clone_lib="NT2RM4"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
                     precursor cells"
BASE COUNT      244 a    149 c    168 g    254 t      4 others
ORIGIN


alignment_scores:
            Quality:  795.00            Length:    147
              Ratio:  5.408               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-613-508-4 x AU125557   ..

 Align seg 1/1  to: AU125557  from: 1  to: 819

     131 AsnGlyProValAspLeuLysLysIleThrAsnPhePheArgGlyAspAr 147
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       3 AATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCG 52

     147 gCysArgSerLeuThrGlyLysProLysLeuPheIleIleGlnAlaCysA 164
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      53 TTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCC 102

     164 rgGlyThrGluLeuAspCysGlyIleGluThrAspSerGlyValAspAsp 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     103 GTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGAT 152

     181 AspMetAlaCysHisLysIleProValGluAlaAspPheLeuTyrAlaTy 197
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     153 GACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATA 202

     197 rSerThrAlaProGlyTyrTyrSerTrpArgAsnSerLysAspGlySerT 214
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     203 CTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCT 252

     214 rpPheIleGlnSerLeuCysAlaMetLeuLysGlnTyrAlaAspLysLeu 230
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     253 GGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTT 302

     231 GluPheMetHisIleLeuThrArgValAsnArgLysValAlaThrGluPh 247
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     303 GAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATT 352

     247 eGluSerPheSerPheAspAlaThrPheHisAlaLysLysGlnIleProC 264
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     353 TGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCAT 402

     264 ysIleValSerMetLeuThrLysGluLeuTyrPheTyrHis 277
         |||||||||||||||||||||||||||||||||||||||
     403 GTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 443

seq_name: gb_est1:AW951022

seq_documentation_block:
LOCUS       AW951022                599 bp    mRNA    linear   EST 01-JUN-2000
DEFINITION  EST363092 MAGE resequences, MAGA Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW951022
VERSION     AW951022.1  GI:8140687
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 599)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
            ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 22
            Seq primer: Reverse.
FEATURES             Location/Qualifiers
     source          1. .599
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGA"
                     /note="Vector: pBluescriptSKm"
BASE COUNT      218 a     85 c    135 g    160 t      1 others
ORIGIN


alignment_scores:
            Quality:  794.50            Length:    173
              Ratio:  4.757               Gaps:      1
 Percent Similarity:  96.532   Percent Identity: 92.486
```

```
alignment_block:
 US-09-613-508-4 x AW951022   ..

 Align seg 1/1  to: AW951022  from: 1  to: 599

       1 MetGluAsnThrGluAsnSerValAspSerLysSerIleLysAsnLeuGl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      38 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGA 87

      17 uProLysIleIleHisGlySerGluSerMetAspSerGlyIleSerLeuA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      88 ACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGG 137

      34 spAsnSerTyrLysMetAspTyrProGluMetGlyLeuCysIleIleIle 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     138 ACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATT 187

      51 AsnAsnLysAsnPheHisLysSerThrGlyMetThrSerArgSerGlyTh 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     188 AATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTAC 237

      67 rAspValAspAlaAlaAsnLeuArgGluThrPheArgAsnLeuLysTyrG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     238 AGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATG 287

      84 luValArgAsnLysAsnAspLeuThrArgGluGluIleValGluLeuMet 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     288 AAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 337

     101 ArgAspValSerLysGluAspHisSerLysArgSerSerPheValCysVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     338 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGT 387

     117 lLeuLeuSerHisGlyGluGluGlyIleIlePheGlyThrAsnGlyProV 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     388 GCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTG 437

     134 alAspLeuLysLysIleThrAsnPhePheArgGlyAspArgCysArgSer 150
         ||||||||||||||||||||||||||||||||||||   :::|||
     438 TTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTG.AAAAGT 486

     151 LeuThrGlyLysProLysLeuPheIleIleGlnAlaCysArgGlyThrGl 167
         |||||||||||||||:::   |||   :::::::::      ||||||
     487 CTAACTGGAAAACCCAAC...TTTTAATTTTCAGGCCTGCCGGGTACANA 533

     167 uLeuAspCysGlyIleGlu 173
         |||||||||:::::|||
     534 ACTTGACTGTGCATTGGAA 552

seq_name: gb_est2:BE284569

seq_documentation_block:
LOCUS       BE284569               705 bp    mRNA    linear   EST 26-OCT-2000
DEFINITION  601087184F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:3482156 5',
            mRNA sequence.
ACCESSION   BE284569
VERSION     BE284569.1  GI:9161595
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 705)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8511  row: g  column: 21
            High quality sequence stop: 574.
FEATURES             Location/Qualifiers
     source          1. .705
                     /organism="Mus musculus"
                     /strain="C57/B6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3482156"
                     /clone_lib="NCI_CGAP_Mam5"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="7 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Robin Humphreys,
                     NIH"
BASE COUNT      207 a    147 c    198 g    153 t
ORIGIN


alignment_scores:
            Quality:  742.00              Length:    187
              Ratio:  4.240                 Gaps:      3
 Percent Similarity:  93.583   Percent Identity: 82.353

alignment_block:
 US-09-613-508-4 x BE284569   ..

 Align seg 1/1  to: BE284569  from: 1  to: 705

      89 AsnAspLeuThrArgGluGluIleValGluLeuMetArgAspValSerLy 105
         ||||||||||||||||||:::|||:::|||||||||   :::||||||||
       3 AATGATCTTACTCGTGAAGACATTTTGGAATTAATGGATAGTGTTTCTAA 52

     105 sGluAspHisSerLysArgSerSerPheValCysValLeuLeuSerHisG 122
         ||||||||||||||||||||||||   |||||||||:::|||||||||||
      53 GGAAGATCATAGCAAAAGGAGCAGCTT.GTGTGTGTGATTCTAAGCCATG 101

     122 lyGluGluGlyIleIlePheGlyThrAsnGlyProValAspLeuLysLys 138
         ||:::||||||:::|||:::|||||||||||||||||||   |||||||||
     102 GTGATGAAGGGGTCATTTATGGGACAAATGGGCCTGTGAA.CTGAAAAAG 150

     139 IleThrAsnPhePheArgGlyAspArgCysArgSerLeuThrGlyLysPr 155
         :::|||:::||||||||||||||||   |||||||||||||||||||||||
     151 TTGACTAGCTTCTTCAGAGGCGACTACTGCCGGAGTCTGACTGGAAAGCC 200

     155 oLysLeuPheIleIleGlnAlaCysArgGlyThrGluLeuAspCysGlyI 172
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 GAAACTCTTCATCATTCAGGCCTGCCGGGGTACGGAGCTGGACTGTGGCA 250

     172 leGluThrAspSerGlyValAspAspAspMetAlaCysHisLysIlePro 188
         ||||||||||||||||||:::|||:::::|||||||||:::|||||||||
     251 TTGAGACAGACAGTGGGACTGATGAGGAGATGGCTTGCCAGAAGATACCG 300

     189 ValGluAlaAspPheLeuTyrAlaTyrSerThrAlaProGlyTyrTyrSe 205
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 GTGGAGGCTGACTTCCTGTATGCTTACTCTACAGCACCTGGTTACTATTC 350

     205 rTrpArgAsnSerLysAspGlySerTrpPheIleGlnSer.LeuCysAla 221
         |||||||||||||||||||||||||||||||||||| |||||||:::
     351 CTGGAGAAATTCAAAGGACGGGTCGTGGTTCATCCAGTCCCTTATGCAGC 400

     222 MetLeuLysGlnTyrAlaAspLysLeuGluPheMetHisIleLeuThrAr 238
         |||||||||   ||||||   ||||||||||:::||||||||||||||||
     401 ATGCTGAAGCTGTACGCGCACAAGCTAGAATATATGCACATTCTCACTCG 450

     238 gValAsnArgLysValAlaThrGluPheGluSerPheSerPheAspAla. 254
         ||||:::|||||||||||||||||||||||||||||||:::|||:::
     451 CGTTACCAGGAAGGTGGCAACGGAATTCGAGTCCTTCTCACTGGACTCCA 500
```

```
255 ThrPheHisAla.LysLysGlnIleProCysIleValSerMetLeuThrL 271
    ||||||||||| |||||      |||         ||||||||||||
501 ACTTTCCACGCAAAAGAAACCAGATCCCCGGTATGGGTCCATGCTCACGA 550

271 ysGlu 272
    ||:::
551 AAGAC 555
```

```
OM of: US-09-613-508-2 to: GenEmbl:*   out_format : pfs

Date: Aug 15, 2002  8:46 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO_spool/US09613508/runat_14082002_112647_7223/app_query.fasta_1.701
 -DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
 -MINLEN=0 -MAXLEN=2000000000 -USER=US09613508_@CGN1_1_4651
 -NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-613-508-2
Query length: 303
Database: GenEmbl:*
Database sequences: 1797656
Database length: 1873333701
Search time (sec): 2733.900000



score_list:
Sequence            Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_pr:HSU39613       + 1610.00 3017.33  1.1e-159  912  ! U39613 Human cysteine protease
gb_pat:AX134417      + 1606.00 3000.78  8.8e-159 2309  ! AX134417 Sequence 13 from Pate
gb_pr:HSU37448       + 1606.00 3000.78  8.8e-159 2309  ! U37448 Human Mch3 isoform alph
gb_pr:HSU67320       + 1606.00 3000.76  8.8e-159 2314  ! U67320 Human Lice2 gamma cyste
gb_pr:HSU40281       + 1606.00 3000.68  8.9e-159 2335  ! U40281 Human cysteine protease
gb_pr:HSU67319       + 1606.00 2999.63  1.0e-158 2602  ! U67319 Human Lice2 beta cystei
gb_pr:HSU67206       + 1606.00 2999.41  1.1e-158 2661  ! U67206 Homo sapiens Lice2 alph
gb_pr:BC015799       + 1602.00 2993.03  2.4e-158 2361  ! BC015799 Homo sapiens, caspase
gb_pat:AR103058      + 1580.00 2956.83  2.5e-156 1371  ! AR103058 Sequence 1 from paten
gb_ro:AF072124       + 1395.00 2610.60  4.8e-137 1063  ! AF072124 Rattus norvegicus cas
gb_ro:MAU47332       + 1394.00 2602.89  1.3e-136 1939  ! U47332 Mesocricetus auratus MC
gb_pr:HSU37449       + 1381.50 2577.35  3.4e-135 2378  ! U37449 Human Mch3 isoform beta
gb_ro:MMU67321       + 1346.00 2512.09  1.5e-131 2006  ! U67321 Mus musculus Lice2 cyst
gb_ro:BC005428       + 1346.00 2510.74  1.7e-131 2305  ! BC005428 Mus musculus, caspase
gb_ro:MMCASP7        + 1346.00 2510.46  1.8e-131 2372  ! Y13088 M.musculus mRNA for cas
gb_ro:D86353         + 1336.00 2500.35  6.6e-131  964  ! D86353 Mouse mRNA for Mch3, com
gb_ov:AB038170       + 1025.00 1909.65  5.3e-98  1534  ! AB038170 Xenopus laevis mRNA f
gb_ov:AB047003       +  744.50 1386.69  7.1e-69  849  ! AB047003 Danio rerio mRNA for c
gb_ro:CGU27463       +  744.00 1375.69  2.9e-68 2395  ! U27463 Cricetinae gen. sp. CPP
gb_ov:AF083029       +  743.00 1378.58  2.0e-68 1464  ! AF083029 Gallus gallus caspase
gb_pat:AX207740      +  727.50 1335.00  5.4e-66 6436  ! AX207740 Sequence 38 from Pate
gb_pat:AX207741      +  727.50 1334.60  5.6e-66 6703  ! AX207741 Sequence 39 from Pate
gb_pat:AX207742      +  727.50 1334.26  5.9e-66 6946  ! AX207742 Sequence 40 from Pate
gb_pat:AX207743      +  727.50 1333.93  6.2e-66 7189  ! AX207743 Sequence 41 from Pate
gb_pat:AX207737      +  721.00 1342.57  2.0e-66  834  ! AX207737 Sequence 35 from Paten
gb_pr:HSU26943       +  721.00 1342.57  2.0e-66  834  ! U26943 Human cysteine protease
gb_pat:AR109589      +  721.00 1340.80  2.6e-66 1001  ! AR109589 Sequence 5 from paten
gb_pr:HSU13738       +  721.00 1340.80  2.6e-66 1001  ! U13738 Human cysteine protease
gb_pat:AR103059      +  721.00 1339.38  3.1e-66 1159  ! AR103059 Sequence 3 from paten
gb_pr:BC016926       +  721.00 1331.35  8.6e-66 2652  ! BC016926 Homo sapiens, clone M
gb_om:AB029345       +  720.50 1341.63  2.3e-66  834  ! AB029345 Sus scrofa mRNA for ca
gb_pr:HSA413269      +  719.00 1338.80  3.3e-66  834  ! AJ413269 Homo sapiens mRNA for
gb_pat:AR172878      +  718.00 1325.75  1.8e-65 2635  ! AR172878 Sequence 3 from paten
gb_pat:AX134409      +  718.00 1325.75  1.8e-65 2635  ! AX134409 Sequence 5 from Paten
gb_pat:AX331135      +  718.00 1325.75  1.8e-65 2635  ! AX331135 Sequence 1644 from Pa
gb_pr:HSU13737       +  718.00 1325.75  1.8e-65 2635  ! U13737 Human cysteine protease
gb_pat:AX207746      +  717.50 1315.00  7.0e-65 7248  ! AX207746 Sequence 44 from Pate
gb_ro:RNU49930       +  716.00 1329.23  1.1e-65 1249  ! U49930 Rattus norvegicus ICE-l
gb_pat:AR109591      +  711.00 1321.95  2.9e-65 1001  ! AR109591 Sequence 9 from paten
gb_ro:RNU84410       +  709.00 1317.43  5.1e-65 1081  ! U84410 Rattus norvegicus inter
gb_pat:AR059697      +  709.00 1316.61  5.7e-65 1177  ! AR059697 Sequence 1 from pa
gb_pat:AR059698      +  709.00 1316.61  5.7e-65 1177  ! AR059698 Sequence 3 from pa
gb_pat:AX134427      +  707.00 1316.75  5.6e-65  786  ! AX134427 Sequence 23 from Pa
gb_ro:D86352         +  703.00 1308.19  1.7e-64  874  ! D86352 Mouse mRNA for CPP32,
gb_pat:AR172885      +  703.00 1304.36  2.7e-64 1297  ! AR172885 Sequence 10 from p

seq_name: gb_pr:HSU39613

seq_documentation_block:
LOCUS       HSU39613                912 bp    mRNA    linear   PRI 19-JAN-1996
DEFINITION  Human cysteine protease ICE-LAP3 mRNA, complete cds.
ACCESSION   U39613
VERSION     U39613.1  GI:1125072
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 912)
  AUTHORS   Duan,H., Chinnaiyan,A.M., Hudson,P.L., Wing,J.P., He,W.W. and
            Dixit,V.M.
  TITLE     ICE-LAP3, a novel mammalian homologue of the Caenorhabditis elegans
            cell death protein Ced-3 is activated during Fas- and tumor
            necrosis factor-induced apoptosis
  JOURNAL   J. Biol. Chem. 271 (3), 1621-1625 (1996)
  MEDLINE   96139498
REFERENCE   2  (bases 1 to 912)
  AUTHORS   Duan,H., Chinnaiyan,A.M., Hudson,P.L., Wing,J.P., He,W.W. and
            Dixit,V.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-OCT-1995) Hangjun Duan, Pathology, University of
            Michigan, 1150 W Medical Center Dr MSRBI Room 7514, Ann Arbor, MI
            48109, USA
FEATURES             Location/Qualifiers
     source          1. .912
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .912
                     /function="cysteine protease"
                     /note="member of ICE/GED-3 gene family of cysteine
                     proteases"
                     /codon_start=1
                     /product="ICE-LAP3"
                     /protein_id="AAC50346.1"
                     /db_xref="GI:1125073"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDAIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      263 a    219 c    223 g    207 t
ORIGIN


alignment_scores:
            Quality: 1610.00              Length:    303
              Ratio:  5.314                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.670

alignment_block:
 US-09-613-508-2 x HSU39613   ..

 Align seg 1/1  to: HSU39613  from: 1  to: 912

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 50

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 100
```

```
 34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
101 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 150

 51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
    ||||||||||||||||||||||||||||||||||||||||||||||||||
151 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 200

 67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
    ||||||||||||||||||||||||||||||||||||||||||||||||||
201 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 250

 84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
    ||||||||||||||||||||||||||||||||||||||||||||||||||
251 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 300

101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
    ||||||||||||||||||||||||||||||||||||||||||||||||||
301 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 350

117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
    ||||||||||||||||||||||||||||||||||||||||||||||||||
351 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 400

134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
401 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 450

151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
    ||||||||||||||||||||||||||||||||||||||||||||||||||
451 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 500

167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
    ||||||||||||||||||||||||||||||||||||||||||||||||||
501 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 550

184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
551 AGGCTTGCCGAGGGACCGAGCTTGATGATGCCATCCAGGCCGACTCGGGG 600

201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
    ||||||||||||||||||||||||||||||||||||||||||||||||||
601 CCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGC 650

217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
    ||||||||||||||||||||||||||||||||||||||||||||||||||
651 TGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGA 700

234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
    ||||||||||||||||||||||||||||||||||||||||||||||||||
701 GCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAG 750

251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
    ||||||||||||:::|||||||||||||||||||||||||||||||||||
751 GAGCACGGAAAAGACCTGGAAATCATGCAGATCCTCACCAGGGTGAATGA 800

267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
    ||||||||||||||||||||||||||||||||||||||||||||||||||
801 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 850

284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
851 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTAC 900

301 PheSerGln 303
    ||||||||
901 TTCAGTCAA 909
```

seq_name: gb_pat:AX134417

seq_documentation_block:

LOCUS AX134417 2309 bp DNA linear PAT 29-MAY-2001
DEFINITION Sequence 13 from Patent WO0129232.

ACCESSION AX134417
VERSION AX134417.1 GI:14271063
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2309)
AUTHORS Cordell,B. and Li,Y.
TITLE Functional cloning of genes encoding proteins/enzymes involved in proteolytic cleavage
JOURNAL Patent: WO 0129232-A 13 26-APR-2001;
Scios Inc. (US)

FEATURES Location/Qualifiers
source 1. .2309
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS 44. .955
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC39525.1"
/db_xref="GI:14271064"
/translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT 680 a 453 c 531 g 645 t
ORIGIN

alignment_scores:
Quality: 1606.00 Length: 303
Ratio: 5.300 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.340

alignment_block:
US-09-613-508-2 x AX134417 ..

Align seg 1/1 to: AX134417 from: 1 to: 2309

```
  1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 44 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 93

 17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 94 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 143

 34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
144 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 193

 51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
    ||||||||||||||||||||||||||||||||||||||||||||||||||
194 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 243

 67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
    ||||||||||||||||||||||||||||||||||||||||||||||||||
244 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 293

 84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
    ||||||||||||||||||||||||||||||||||||||||||||||||||
294 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 343

101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
    ||||||||||||||||||||||||||||||||||||||||||||||||||
344 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 393

117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
    ||||||||||||||||||||||||||||||||||||||||||||||||||
394 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 443
```

```
    134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    444 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 493

    151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    494 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 543

    167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    544 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 593

    184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
        ||||||||||||||||||||||||||||:::||||||||||||||||||
    594 AGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGG 643

    201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    644 CCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGC 693

    217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    694 TGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGA 743

    234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    744 GCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAG 793

    251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
        ||||||||||||:::||||||||||||||||||||||||||||||||||
    794 GAGCACGGAAAAGACCTGGAAATCATGCAGATCCTCACCAGGGTGAATGA 843

    267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    844 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 893

    284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    894 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTAC 943

    301 PheSerGln 303
        ||||||||
    944 TTCAGTCAA 952

seq_name: gb_pr:HSU37448

seq_documentation_block:
LOCUS       HSU37448                2309 bp    mRNA    linear   PRI 13-DEC-1995
DEFINITION  Human Mch3 isoform alpha (Mch3) mRNA, complete cds.
ACCESSION   U37448
VERSION     U37448.1  GI:1117846
KEYWORDS    Mch3 isoform alpha.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2309)
  AUTHORS   Fernandes-Alnemri,T., Litwack,G. and Alnemri,E.S.
  TITLE     CPP32, a novel human apoptotic protein with homology to
            Caenorhabditis elegans cell death protein Ced-3 and mammalian
            interleukin-1 beta-converting enzyme
  JOURNAL   J. Biol. Chem. 269 (49), 30761-30764 (1994)
  MEDLINE   95074098
REFERENCE   2  (bases 1 to 2309)
  AUTHORS   Fernandes-Alnemri,T., Takahashi,A., Armstrong,R., Krebs,J.,
            Fritz,L., Tomaselli,K.J., Wang,L., Yu,Z., Croce,C.M., Salveson,G.,
            Earnshaw,W.C., Litwack,G. and Alnemri,E.S.
  TITLE     Mch3, a novel human apoptotic cysteine protease highly related to
            CPP32
  JOURNAL   Cancer Res. 55 (24), 6045-6052 (1995)
  MEDLINE   96105019
REFERENCE   3  (bases 1 to 2309)
  AUTHORS   Alnemri,E.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-1995) Emad S. Alnemri, Pharmacology, Thomas
            Jefferson University, Jefferson Cancer Institute, 233, S. Tenth
            Street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .2309
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="T50-8A.1"
                     /cell_line="Jurkat"
                     /cell_type="T-lymphocyte"
     gene            1. .2309
                     /gene="Mch3"
     CDS             44. .955
                     /gene="Mch3"
                     /function="cysteine protease"
                     /note="new member of ICE-like cysteine protease family"
                     /codon_start=1
                     /evidence=experimental
                     /product="Mch3 isoform alpha"
                     /protein_id="AAC50303.1"
                     /db_xref="GI:1117847"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
     polyA_site      2309
                     /gene="Mch3"
BASE COUNT      680 a    453 c    531 g    645 t
ORIGIN


alignment_scores:
            Quality: 1606.00            Length:    303
              Ratio:  5.300              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.340

alignment_block:
 US-09-613-508-2 x HSU37448   ..

 Align seg 1/1  to: HSU37448  from: 1  to: 2309

      1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     44 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 93

     17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     94 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 143

     34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    144 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 193

     51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    194 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 243

     67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    244 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 293

     84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    294 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 343

    101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    344 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 393
```

```
    117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    394 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 443

    134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    444 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 493

    151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    494 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 543

    167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    544 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 593

    184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
        ||||||||||||||||||||||||||||||:::|||||||||||||||||
    594 AGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGG 643

    201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    644 CCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGC 693

    217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    694 TGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGA 743

    234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    744 GCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAG 793

    251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
        ||||||||||||:::||||||||||||||||||||||||||||||||||||
    794 GAGCACGGAAAAGACCTGGAAATCATGCAGATCCTCACCAGGGTGAATGA 843

    267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    844 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 893

    284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    894 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTAC 943

    301 PheSerGln 303
        ||||||||
    944 TTCAGTCAA 952

seq_name: gb_pr:HSU67320

seq_documentation_block:
LOCUS       HSU67320              2314 bp    mRNA    linear   PRI 19-MAR-1997
DEFINITION  Human Lice2 gamma cysteine protease mRNA, complete cds.
ACCESSION   U67320
VERSION     U67320.1  GI:1894914
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2314)
  AUTHORS   Juan,T.S., McNiece,I.K., Argento,J.M., Jenkins,N.A., Gilbert,D.J.,
            Copeland,N.G. and Fletcher,F.A.
  TITLE     Identification and mapping of Casp7, a cysteine protease resembling
            CPP32 beta, interleukin-1 beta converting enzyme, and CED-3
  JOURNAL   Genomics 40 (1), 86-93 (1997)
  MEDLINE   97224489
REFERENCE   2  (bases 1 to 2314)
  AUTHORS   Juan,T.S.C., McNiece,I.K., Argento,J.M., Jenkins,N.A.,
            Gilbert,D.J., Copeland,N.G. and Fletcher,F.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-1996) Developmental Hematology, Amgen, Inc., 1840
            DeHavilland Dr., Thousand Oaks, CA 91320, USA

FEATURES             Location/Qualifiers
     source          1. .2314
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             40. .951
                     /note="similar to ICE and CED-3 cysteine protease"
                     /codon_start=1
                     /product="Lice2 gamma cysteine protease"
                     /protein_id="AAC51153.1"
                     /db_xref="GI:1894915"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      680 a    459 c    526 g    648 t      1 others
ORIGIN


alignment_scores:
            Quality: 1606.00              Length:    303
              Ratio:  5.300                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.340

alignment_block:
 US-09-613-508-2 x HSU67320   ..

 Align seg 1/1  to: HSU67320  from: 1  to: 2314

      1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
     40 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 89

     17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
     90 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 139

     34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    140 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 189

     51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    190 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 239

     67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    240 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 289

     84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    290 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 339

    101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    340 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 389

    117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    390 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 439

    134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    440 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 489

    151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    490 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 539

    167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    540 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 589
```

```
     184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
         ||||||||||||||||||||||||||||||:::||||||||||||||||||
     590 AGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGG 639

     201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     640 CCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGC 689

     217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     690 TGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCATGGAGGA 739

     234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     740 GCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAG 789

     251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
         ||||||||||||:::||||||||||||||||||||||||||||||||||
     790 GAGCACGGAAAAGACCTGGAAATCATGCAGATCCTCACCAGGGTGAATGA 839

     267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     840 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 889

     284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     890 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTAC 939

     301 PheSerGln 303
         ||||||||
     940 TTCAGTCAA 948

seq_name: gb_pr:HSU40281

seq_documentation_block:
LOCUS       HSU40281                2335 bp    mRNA    linear   PRI 27-JAN-1996
DEFINITION  Human cysteine protease CMH-1 mRNA, complete cds.
ACCESSION   U40281
VERSION     U40281.1  GI:1167554
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2335)
  AUTHORS   Lippke,J.A., Gu,Y., Sarnecki,C., Caron,P.R. and Su,M.S.
  TITLE     Identification and characterization of CPP32/Mch2 homolog 1, a
            novel cysteine protease similar to CPP32
  JOURNAL   J. Biol. Chem. 271 (4), 1825-1828 (1996)
  MEDLINE   96147144
REFERENCE   2  (bases 1 to 2335)
  AUTHORS   Caron,P.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-NOV-1995) Paul R. Caron, Vertex Pharmaceuticals, 40
            Allston St., Cambridge, MA 02139, USA
FEATURES             Location/Qualifiers
     source          1. .2335
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="IMAGE consortium clone 72778"
     CDS             48. .959
                     /function="cysteine protease"
                     /note="similar to CPP32 and to MCH2; related to Ced3 and
                     interleukin-1 beta converting enzyme"
                     /codon_start=1
                     /product="CMH-1"
                     /protein_id="AAC50352.1"
                     /db_xref="GI:1167555"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      700 a    454 c    535 g    646 t
ORIGIN


alignment_scores:
            Quality: 1606.00              Length:    303
              Ratio:  5.300                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.340

alignment_block:
 US-09-613-508-2 x HSU40281   ..

 Align seg 1/1  to: HSU40281  from: 1  to: 2335

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      48 ATGGCAGATGATCAGGGCTGTATTGAGGAGCAGGGGGTTGAGGATTCAGC 97

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      98 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 147

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     148 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 197

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     198 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 247

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     248 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 297

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     298 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 347

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     348 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 397

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     398 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 447

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     448 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 497

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     498 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 547

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     548 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 597

     184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
         ||||||||||||||||||||||||||||||:::||||||||||||||||||
     598 AGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGG 647

     201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     648 CCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGC 697

     217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     698 TGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGA 747

     234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      748 GCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAG 797

      251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
          ||||||||||||:::||||||||||||||||||||||||||||||||||||
      798 GAGCACGGAAAAGACCTGGAAATCATGCAGATCCTCACCAGGGTGAATGA 847

      267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      848 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 897

      284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      898 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTAC 947

      301 PheSerGln 303
          ||||||||
      948 TTCAGTCAA 956

seq_name: gb_pr:HSU67319

seq_documentation_block:
LOCUS       HSU67319               2602 bp    mRNA    linear   PRI 19-MAR-1997
DEFINITION  Human Lice2 beta cysteine protease mRNA, complete cds.
ACCESSION   U67319
VERSION     U67319.1  GI:1894912
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2602)
  AUTHORS   Juan,T.S., McNiece,I.K., Argento,J.M., Jenkins,N.A., Gilbert,D.J.,
            Copeland,N.G. and Fletcher,F.A.
  TITLE     Identification and mapping of Casp7, a cysteine protease resembling
            CPP32 beta, interleukin-1 beta converting enzyme, and CED-3
  JOURNAL   Genomics 40 (1), 86-93 (1997)
  MEDLINE   97224489
REFERENCE   2  (bases 1 to 2602)
  AUTHORS   Juan,T.S.C., McNiece,I.K., Argento,J.M., Jenkins,N.A.,
            Gilbert,D.J., Copeland,N.G. and Fletcher,F.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-1996) Developmental Hematology, Amgen, Inc., 1840
            DeHavilland Dr., Thousand Oaks, CA 91320, USA
FEATURES             Location/Qualifiers
     source          1. .2602
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             229. .1239
                     /note="similar to ICE and CED-3 cysteine protease"
                     /codon_start=1
                     /product="Lice2 beta cysteine protease"
                     /protein_id="AAC51152.1"
                     /db_xref="GI:1894913"
                     /translation="MDCVGWPPGRKWHLEKNTSCGGSSGICASYVTQMADDQGCIEEQ
                     GVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVTMRSIKTTRDRVPTYQYNMNFEKLGK
                     CIIINNKNFDKVTGMGVRNGTDKDAEALFKCFRSLGFDVIVYNDCSCAKMQDLLKKAS
                     EEDHTNAACFACILLSHGEENVIYGKDGVTPIKDLTAHFRGDRCKTLLEKPKLFFIQA
                     CRGTELDDGIQADSGPINDTDANPRYKIPVEADFLFAYSTVPGYYSWRSPGRGSWFVQ
                     ALCSILEEHGKDLEIMQILTRVNDRVARHFESQSDDPHFHEKKQIPCVVSMLTKELYF
                     SQ"
BASE COUNT      768 a    515 c    616 g    702 t      1 others
ORIGIN


alignment_scores:
            Quality: 1606.00             Length:    303
              Ratio:  5.300               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.340

alignment_block:
 US-09-613-508-2 x HSU67319   ..
```

```
Align seg 1/1  to: HSU67319  from: 1  to: 2602

        1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      328 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 377

       17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      378 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 427

       34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      428 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 477

       51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      478 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 527

       67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      528 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 577

       84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      578 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 627

      101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      628 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 677

      117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      678 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 727

      134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      728 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 777

      151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      778 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 827

      167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      828 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 877

      184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
          ||||||||||||||||||||||||||||||:::||||||||||||||||||
      878 AGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGG 927

      201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      928 CCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGC 977

      217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      978 TGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCATGGAGGA 1027

      234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1028 GCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAG 1077

      251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
          ||||||||||||:::||||||||||||||||||||||||||||||||||||
     1078 GAGCACGGAAAAGACCTGGAAATCATGCAGATCCTCACCAGGGTGAATGA 1127

      267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1128 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 1177

      284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1178 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTAC 1227
```

```
     301 PheSerGln 303
         |||||||||
    1228 TTCAGTCAA 1236

seq_name: gb_pr:HSU67206

seq_documentation_block:
LOCUS       HSU67206               2661 bp    mRNA    linear   PRI 03-JAN-2000
DEFINITION  Homo sapiens Lice2 alpha mRNA, complete cds.
ACCESSION   U67206
VERSION     U67206.1  GI:6649572
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2661)
  AUTHORS   Juan,T.S., McNiece,I.K., Argento,J.M., Jenkins,N.A., Gilbert,D.J.,
            Copeland,N.G. and Fletcher,F.A.
  TITLE     Identification and mapping of Casp7, a cysteine protease resembling
            CPP32 beta, interleukin-1 beta converting enzyme, and CED-3
  JOURNAL   Genomics 40 (1), 86-93 (1997)
  MEDLINE   97224489
   PUBMED   9070923
REFERENCE   2  (bases 1 to 2661)
  AUTHORS   Juan,T.S.C., McNiece,I.K., Argento,J.M., Jenkins,N.A.,
            Gilbert,D.J., Copeland,N.G. and Fletcher,F.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-AUG-1996) Developmental Hematology, Amgen, Inc., 1840
            DeHavilland Dr., Thousand Oaks, CA 91320, USA
FEATURES             Location/Qualifiers
     source          1. .2661
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             387. .1298
                     /note="like-ICE cysteine protease"
                     /codon_start=1
                     /product="Lice2 alpha"
                     /protein_id="AAF21460.1"
                     /db_xref="GI:6649573"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      740 a    569 c    642 g    709 t      1 others
ORIGIN


alignment_scores:
           Quality: 1606.00            Length:    303
             Ratio:  5.300              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.340

alignment_block:
 US-09-613-508-2 x HSU67206   ..

 Align seg 1/1  to: HSU67206  from: 1  to: 2661

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     387 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 436

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     437 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 486

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     487 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 536

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     537 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 586

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     587 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 636

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     637 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 686

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     687 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 736

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     737 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 786

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     787 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 836

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     837 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 886

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     887 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 936

     184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
         ||||||||||||||||||||||||||||||:::||||||||||||||||||
     937 AGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGG 986

     201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     987 CCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGC 1036

     217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1037 TGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCATGGAGGA 1086

     234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1087 GCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAG 1136

     251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
         ||||||||||||:::||||||||||||||||||||||||||||||||||
    1137 GAGCACGGAAAAGACCTGGAAATCATGCAGATCCTCACCAGGGTGAATGA 1186

     267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1187 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 1236

     284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1237 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTAC 1286

     301 PheSerGln 303
         ||||||||
    1287 TTCAGTCAA 1295

seq_name: gb_pr:BC015799

seq_documentation_block:
LOCUS       BC015799               2361 bp    mRNA    linear   PRI 29-OCT-2001
DEFINITION  Homo sapiens, caspase 7, apoptosis-related cysteine protease, clone
            MGC:9176 IMAGE:3921455, mRNA, complete cds.
ACCESSION   BC015799
VERSION     BC015799.1  GI:16041818
KEYWORDS    MGC.
```

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2361)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-OCT-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 15 Row: n Column: 12
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 15718696.
FEATURES              Location/Qualifiers
     source           1. .2361
                      /organism="Homo sapiens"
                      /db_xref="LocusID:840"
                      /db_xref="taxon:9606"
                      /clone="MGC:9176 IMAGE:3921455"
                      /tissue_type="Skin, melanotic melanoma."
                      /clone_lib="NIH_MGC_72"
                      /lab_host="DH10B"
                      /note="Vector: pCMV-SPORT6"
     CDS              79. .990
                      /codon_start=1
                      /product="caspase 7, apoptosis-related cysteine protease"
                      /protein_id="AAH15799.1"
                      /db_xref="GI:16041819"
                      /translation="MADEQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                      MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                      SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                      DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                      FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                      SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      704 a    464 c    545 g    648 t
ORIGIN


alignment_scores:
           Quality: 1602.00            Length:    303
             Ratio:  5.287               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.010

alignment_block:
 US-09-613-508-2 x BC015799   ..

 Align seg 1/1  to: BC015799  from: 1  to: 2361

      1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
        |||||||||::::|||||||||||||||||||||||||||||||||||||
     79 ATGGCAGATGAGCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 128

     17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    129 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 178

     34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    179 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 228

     51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    229 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 278

     67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    279 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 328

     84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    329 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 378

    101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    379 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 428

    117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    429 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 478

    134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    479 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 528

    151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    529 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 578

    167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    579 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 628

    184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
        ||||||||||||||||||||||||||||||::::|||||||||||||||||
    629 AGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGG 678

    201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    679 CCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGC 728

    217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    729 TGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGA 778

    234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    779 GCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAG 828

    251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
        ||||||||||||:::||||||||||||||||||||||||||||||||||
    829 GAGCACGGAAAAGACCTGGAAATCATGCAGATCCTCACCAGGGTGAATGA 878

    267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    879 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 928

    284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    929 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTAC 978

    301 PheSerGln 303
        ||||||||
    979 TTCAGTCAA 987

seq_name: gb_pat:AR103058
```

```
seq_documentation_block:
LOCUS       AR103058               1371 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 1 from patent US 6087150.
ACCESSION   AR103058
VERSION     AR103058.1  GI:12814646
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1371)
  AUTHORS   He,W.W., Rosen,C.A., Hudson,P.L. and Hastings,G.A.
  TITLE     Interleukin-1 .beta. converting enzyme like apoptosis protease-3
            and 4
  JOURNAL   Patent: US 6087150-A 1 11-JUL-2000;
FEATURES             Location/Qualifiers
     source          1. .1371
                     /organism="unknown"
BASE COUNT      355 a    357 c    354 g    305 t
ORIGIN


alignment_scores:
            Quality: 1580.00             Length:    305
              Ratio:  5.215               Gaps:      2
 Percent Similarity:  99.344   Percent Identity: 98.689

alignment_block:
 US-09-613-508-2 x AR103058   ..

 Align seg 1/1  to: AR103058  from: 1  to: 1371

      1 MetAlaAspAsp.GlnGlyCysIleGluGluGlnGlyValGluAspSerA 17
        ||||||||||| |||||||||||||||||||||||||||||||||||||||
    339 ATGGCAGATGATTCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAG 388

     17 laAsnGluAspSerValAspAlaLysProAspArgSerSerPheValPro 33
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    389 CAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCG 438

     34 SerLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysTh 50
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    439 TCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGAC 488

     50 rThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysL 67
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    489 CACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGC 538

     67 euGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGly 83
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    539 TGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGT 588

     84 MetGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCy 100
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    589 ATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTG 638

    100 sPheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysA 117
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    639 CTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTG 688

    117 laLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsn 133
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    689 CCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAAT 738

    134 AlaAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIl 150
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    739 GCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAAT 788

    150 eTyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheA 167
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    789 TTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTA 838

    167 rgGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIle 183
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    839 GGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATT 888

    184 GlnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGl 200
        |||||||||||||||||||||||||||||:::|||||||||||||||||||
    889 CAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGG 938

    200 yProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluA 217
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    939 GCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAG 988

    217 laAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArg 233
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    989 CTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGG 1038

    234 SerProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGl 250
        |||||||||||||||||||||||||||||||||||||||||||||||||||
   1039 AGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGA 1088

    250 uGluHisGlyLysGluLeuGluIleMetGlnIleLeu.ThrArgValAsn 266
        |||||||||||||:::|||||||||||||||||||| ||||||||||||
   1089 GGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAAT 1138

    267 AspArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHi 283
        |||||||||||||||||||||||||||||||||||||||||||||||||||
   1139 GACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCA 1188

    283 sGluLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuT 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||
   1189 TGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCT 1238

    300 yrPheSerGln 303
        ||||||||||
   1239 ACTTCAGTCAA 1249

seq_name: gb_ro:AF072124

seq_documentation_block:
LOCUS       AF072124               1063 bp    mRNA    linear   ROD 24-MAY-2000
DEFINITION  Rattus norvegicus caspase-7 mRNA, complete cds.
ACCESSION   AF072124
VERSION     AF072124.1  GI:3243275
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1063)
  AUTHORS   Forghani,F. and Roy,S.
  TITLE     Rat caspase-7 sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1063)
  AUTHORS   Forghani,F. and Roy,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-1998) Biochemistry, Merck Frosst Centre for
            Therapeutic Research, P.O. Box 1005, Pointe-Claire-Dorval, QUE H9H
            3L1, Canada
FEATURES             Location/Qualifiers
     source          1. .1063
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /tissue_type="spleen"
     CDS             3. .914
                     /codon_start=1
                     /product="caspase-7"
                     /protein_id="AAC24011.1"
                     /db_xref="GI:3243276"
                     /translation="MTDDQDCAAELEMADSSTEDGVDAKPDRSTIISSLLWKKKKNAS
                     MCPVSTTRDRVPTYLYRMDFEKMGKCIIINNKNFDKATGMDVRNGTDKDAEALFKCFR
                     SLGFEVTVYNDCSCAKMQDLLRRASEEDHSNSACFACVLLSHGEENLIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRNPGKGSWFVQALCSILNEHGKDLEIMQILTRVNDRVARHFESQ
```

```
                    SDDPRFNEKKQIPCMVSMLTKELYFGR"
BASE COUNT      278 a    268 c    275 g    242 t
ORIGIN


alignment_scores:
            Quality: 1395.00              Length:    303
              Ratio:  4.810                Gaps:      0
 Percent Similarity:  95.710   Percent Identity: 84.488

alignment_block:
 US-09-613-508-2 x AF072124    ..

 Align seg 1/1  to: AF072124  from: 1  to: 1063

       1 MetAlaAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         |||:::|||||||||   |||      |||       :::   ||||||::
       3 ATGACAGACGATCAGGACTGTGCTGCAGAGCTGGAAATGGCAGATTCTTC 52

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ::::||||||:::|||||||||||||||||||||||||:::::::::   |
      53 GACGGAAGACGGAGTTGACGCTAAGCCAGACCGCTCCACCATCATCTCAT 102

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         |||||:::   |||||||||||||||||:::::|||      ::::::|||
     103 CCCTTCTCTGGAAGAAGAAGAAAAATGCCTCTATGTGCCCCGTCAGTACC 152

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||   |||:::|||:::||||||||::
     153 ACCCGGGACCGCGTGCCCACTTACCTGTACCGTATGGACTTCGAGAAGAT 202

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         :|||||||||||||||||||||||||||||||||||||||:::|||||||
     203 GGGCAAATGTATCATCATAAATAACAAGAACTTCGACAAAGCGACAGGTA 252

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||   ||||||||||||||||||||||||||||||||||||||||||||||
     253 TGGACGTCCGGAATGGGACGGACAAAGACGCAGAGGCCCTCTTCAAGTGC 302

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         |||||||||||||||||:::|||   |||||||||||||||||||||||
     303 TTCCGAAGCCTGGGTTTTGAAGTGACTGTCTATAATGACTGCTCTTGCGC 352

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         |||||||||||||||||||:::::|||||||||||||||||||:::|||:
     353 AAAGATGCAGGATCTGCTTAGACGAGCCTCTGAAGAGGACCACAGCAACT 402

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         ::|||||||||||||||:::|||||||||||||||||||||||||:::|||
     403 CAGCCTGCTTCGCCTGCGTCCTGCTGAGCCACGGAGAAGAGAATCTGATT 452

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     453 TATGGGAAAGACGGCGTGACACCGATAAAGGATCTGACAGCTCATTTTAG 502

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     503 GGGAGACCGATGCAAAACCCTGCTAGAGAAACCCAAGCTCTTCTTCATCC 552

     184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
         ||||||||||||||||||||||||||||:::|||||||||||||||||||
     553 AGGCGTGCCGAGGGACGGAGCTGGATGACGGGATCCAGGCTGACTCGGGG 602

     201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     603 CCTATCAACGACACCGACGCTAATCCCCGCTACAAGATCCCGGTGGAAGC 652

     217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||:
     653 TGACTTTCTCTTTGCTTACTCCACGGTTCCAGGCTATTATTCGTGGAGGA 702

     234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
         ::||||||:::|||||||||||||||||||||||||||||||||||||:::
     703 ACCCAGGAAAGGGCTCCTGGTTTGTGCAGGCCCTCTGCTCCATCCTGAAT 752

     251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
         ||||||||||||:::|||||||||||||||||||||||||||||||||||
     753 GAGCATGGCAAGGACCTGGAGATCATGCAGATCCTGACCAGGGTGAACGA 802

     267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
         ||||||||||||||||||||||||||||||||||||||||||:::|||:::|
     803 CAGGGTGGCCAGACACTTCGAGTCCCAGTCTGATGACCCCCGCTTCAATG 852

     284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
         ||||||||||||||||||||:::||||||||||||||||||||||||||||
     853 AGAAGAAGCAGATTCCATGTATGGTGTCCATGCTCACCAAAGAGCTCTAC 902

     301 PheSerGln 303
         |||:::::::
     903 TTTGGCCGT 911

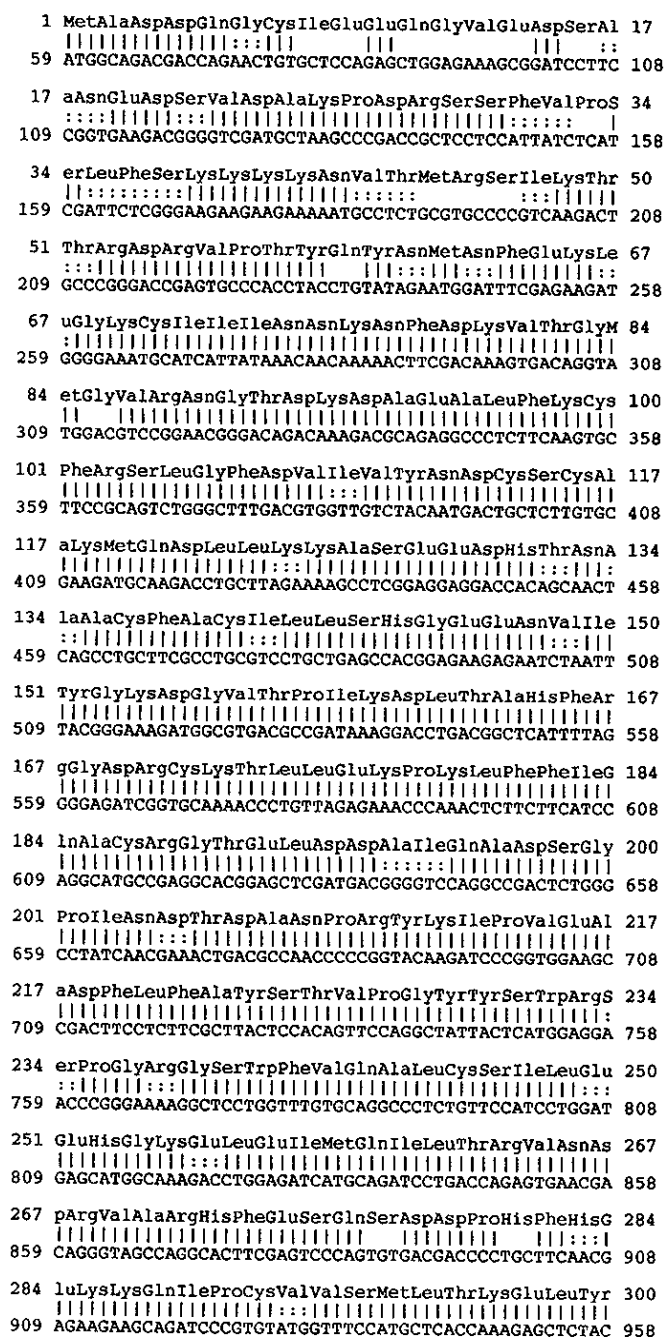
seq_name: gb_ro:MAU47332

seq_documentation_block:
LOCUS       MAU47332                1939 bp    mRNA    linear   ROD 05-JUN-1996
DEFINITION  Mesocricetus auratus MCH3/SCA-2 mRNA, complete cds.
ACCESSION   U47332
VERSION     U47332.1  GI:1245143
KEYWORDS    .
SOURCE      golden hamster.
  ORGANISM  Mesocricetus auratus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Mesocricetus.
REFERENCE   1  (bases 1 to 1939)
  AUTHORS   Pai,J.T., Brown,M.S. and Goldstein,J.L.
  TITLE     Purification and cDNA cloning of a second apoptosis-related
            cysteine protease that cleaves and activates sterol regulatory
            element binding proteins
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 93 (11), 5437-5442 (1996)
  MEDLINE   96224303
REFERENCE   2  (bases 1 to 1939)
  AUTHORS   Pai,J.-T., Brown,M.S. and Goldstein,J.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-1996) Jih-tung Pai, Molecular Genetics,
            University of Texas, Southwestern Medical Center, 5323 Harry Hines
            Blvd., Dallas, TX 75235-9046, USA
FEATURES             Location/Qualifiers
     source          1. .1939
                     /organism="Mesocricetus auratus"
                     /db_xref="taxon:10036"
     CDS             59. .970
                     /note="cysteine protease"
                     /codon_start=1
                     /product="MCH3/SCA-2"
                     /protein_id="AAC52595.1"
                     /db_xref="GI:1245144"
                     /translation="MADDQNCAPELEKADPSGEDGVDAKPDRSSIISSILGKKKKNAS
                     ACPVKTARDRVPTYLYRMDFEKMGKCIIINNKNFDKVTGMDVRNGTDKDAEALFKCFR
                     SLGFDVVVYNDCSCAKMQDLLRKASEEDHSNSACFACVLLSHGEENLIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGVQADSGPINETDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRNPGKGSWFVQALCSILDEHGKDLEIMQILTRVNDRVARHFESQ
                     CDDPCFNEKKQIPCMVSMLTKELYFGR"
BASE COUNT      493 a    464 c    523 g    459 t
ORIGIN


alignment_scores:
            Quality: 1394.00              Length:    303
              Ratio:  4.840                Gaps:      0
 Percent Similarity:  95.050   Percent Identity: 84.158

alignment_block:
 US-09-613-508-2 x MAU47332    ..

 Align seg 1/1  to: MAU47332  from: 1  to: 1939
```




```
     301 PheSerGln 303
         |||::::::
     959 TTCGGCCGT 967

seq_name: gb_pr:HSU37449

seq_documentation_block:
LOCUS       HSU37449                2378 bp    mRNA    linear   PRI 13-DEC-1995
DEFINITION  Human Mch3 isoform beta (Mch3) mRNA, complete cds.
ACCESSION   U37449
VERSION     U37449.1  GI:1117848
KEYWORDS    Mch3 isoform beta.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2378)
  AUTHORS   Fernandes-Alnemri,T., Litwack,G. and Alnemri,E.S.
  TITLE     CPP32, a novel human apoptotic protein with homology to
            Caenorhabditis elegans cell death protein Ced-3 and mammalian
            interleukin-1 beta-converting enzyme
  JOURNAL   J. Biol. Chem. 269 (49), 30761-30764 (1994)
  MEDLINE   95074098
REFERENCE   2  (bases 1 to 2378)
  AUTHORS   Fernandes-Alnemri,T., Takahashi,A., Armstrong,R., Krebs,J.,
            Fritz,L., Tomaselli,K.J., Wang,L., Yu,Z., Croce,C.M., Salveson,G.,
            Earnshaw,W.C., Litwack,G. and Alnemri,E.S.
  TITLE     Mch3, a novel human apoptotic cysteine protease highly related to
            CPP32
  JOURNAL   Cancer Res. 55 (24), 6045-6052 (1995)
  MEDLINE   96105019
REFERENCE   3  (bases 1 to 2378)
  AUTHORS   Alnemri,E.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-1995) Emad S Alnemri, Pharmacology, Thomas
            Jefferson University, Jefferson Cancer Institute, 233, S. Tenth
            Street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .2378
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="T50-6A.1"
                     /cell_line="Jurkat"
                     /cell_type="T-lymphocyte"
     gene            1. .2378
                     /gene="Mch3"
     CDS             147. .908
                     /gene="Mch3"
                     /note="does not have protease activity"
                     /codon_start=1
                     /evidence=experimental
                     /product="Mch3 isoform beta"
                     /protein_id="AAC50304.1"
                     /db_xref="GI:1117849"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENMESCSVTQAGVQ
                     RRDLGRLQPPPPRLAEGPSLMMASRPTRGPSMTQMLILDTRSQWKLTSSSPIPRFQAI
                     TRGGAQEEAPGLCKPSAPSWRSTEKTWKSCRSSPG"
     polyA_site      2378
                     /gene="Mch3"
BASE COUNT      680 a    485 c    566 g    647 t
ORIGIN


alignment_scores:
            Quality: 1381.50              Length:    305
              Ratio:  4.934                Gaps:        3
 Percent Similarity:  91.803   Percent Identity: 88.525

alignment_block:
 US-09-613-508-2 x HSU37449   ..
```

```
Align seg 1/1  to: HSU37449  from: 1  to: 2378

      1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
    147 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 196

     17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
    197 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 246

     34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
    247 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 296

     51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
    297 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 346

     67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
    347 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 396

     84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
    397 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 446

    101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
    447 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 496

    117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
    497 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 546

    134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
    547 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAAT...... 590

    151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
    591 .....ATGGAATCTTGCTCTGTCACCCAGGC 616

    167 gGlyAspArgCysLysThrLeu......LeuGluLysProLysLeuPheP 182
    617 TGGAGTGCAGCGGCGTGATCTCGGAAGACTGCAACCTCCACCTCC..... 661

    182 heIleGlnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAsp 198
    662 .....CAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGAC 706

    199 SerGlyProIleAsnAspThrAspAlaAsnProArgTyrLysIleProVa 215
    707 TCGGGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGT 756

    215 lGluAlaAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerT 232
    757 GGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGT 806

    232 rpArgSerProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIle 248
    807 GGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATC 856

    249 LeuGluGluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgVa 265
    857 CTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCTCACCAGGGT 906

    265 lAsnAspArgValAlaArgHisPheGluSerGlnSerAspAspProHisP 282
    907 GAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACT 956

    282 heHisGluLysLysGlnIleProCysValValSerMetLeuThrLysGlu 298
    957 TCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAA 1006

    299 LeuTyrPheSerGln 303
   1007 CTCTACTTCAGTCAA 1021

seq_name: gb_ro:MMU67321

seq_documentation_block:
LOCUS       MMU67321              2006 bp    mRNA    linear   ROD 19-MAR-1997
DEFINITION  Mus musculus Lice2 cysteine protease (Lice2) mRNA, complete cds.
ACCESSION   U67321
VERSION     U67321.1  GI:1894916
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2006)
  AUTHORS   Juan,T.S., McNiece,I.K., Argento,J.M., Jenkins,N.A., Gilbert,D.J.,
            Copeland,N.G. and Fletcher,F.A.
  TITLE     Identification and mapping of Casp7, a cysteine protease resembling
            CPP32 beta, interleukin-1 beta converting enzyme, and CED-3
  JOURNAL   Genomics 40 (1), 86-93 (1997)
  MEDLINE   97224489
REFERENCE   2  (bases 1 to 2006)
  AUTHORS   Juan,T.S.C., McNiece,I.K., Argento,J.M., Jenkins,N.A.,
            Gilbert,D.J., Copeland,N.G. and Fletcher,F.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-1996) Developmental Hematology, Amgen, Inc., 1840
            DeHavilland Dr., Thousand Oaks, CA 91320, USA
FEATURES             Location/Qualifiers
     source          1. .2006
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="19"
                     /map="Between Adra2a and Aop1"
     gene            1. .2006
                     /gene="Lice2"
     CDS             474. .1496
                     /gene="Lice2"
                     /note="similar to ICE and CED-3 cysteine protease"
                     /codon_start=1
                     /product="Lice2 cysteine protease"
                     /protein_id="AAC53068.1"
                     /db_xref="GI:1894917"
                     /translation="MFTAQVFSESFTKTELLPSTLAEDGRCRGLLAAAVGTMTDDQDC
                     AAELEKVDSSSEDGVDAKPDRSSIISSILLKKKRNASAGPVRTGRDRVPTYLYRMDFQ
                     KMGKCIIINNKNFDKATGMDVRNGTDKDAGALFKCFQNLGFEVTVHNDCSCAKMQDLL
                     RKASEEDHSNSACFACVLLSHGEEDLIYGKDGVTPIKDLTAHFRGDRCKTLLEKPKLF
                     FIQACRGTELDDGIQADSGPINDIDANPRNKIPVEADFLFAYSTVPGYYSWRNPGKGS
                     WFVQALCSILNEHGKDLEIMQILTRVNDRVARHFESQSDDPRFNEKKQIPCMVSMLTK
                     ELYFSR"
BASE COUNT      543 a    490 c    462 g    511 t
ORIGIN


alignment_scores:
           Quality: 1346.00            Length:    303
             Ratio:  4.739              Gaps:      0
 Percent Similarity:  93.729   Percent Identity: 81.518

alignment_block:
 US-09-613-508-2 x MMU67321   ..

 Align seg 1/1  to: MMU67321  from: 1  to: 2006

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
     585 ATGACCGATGATCAGGACTGTGCTGCGGAGCTGGAAAAGGTGGATTCTTC 634

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
     635 CAGCGAAGACGGAGTTGACGCCAAGCCAGACCGCTCCTCTATCATCTCCT 684
```

```
     34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
        ||:::::   |||||||||:::|||:::::      :::::|||
    685 CTATTCTCTTGAAGAAGAAGAGAAATGCCTCTGCGGGCCCCGTCAGGACC 734

     51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
        ||||||||||||||||||||   |||:::|||:::|||:::|||::
    735 GGCCGGGACCGAGTGCCCACTTATCTGTACCGCATGGATTTCCAGAAGAT 784

     67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
        :|||||||||||||||||||||||||||||||||||||:::|||||||
    785 GGGTAAATGCATCATCATAAACAACAAGAACTTCGACAAAGCGACAGGTA 834

     84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
        ||   |||||||||||||||||||||||||   ||||||||||||||||
    835 TGGACGTCCGGAATGGGACGGACAAAGATGCAGGGGCCCTCTTCAAGTGC 884

    101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
        |||:::::|||||||||:::|||   |||:::|||||||||||||||||
    885 TTCCAAAACCTGGGTTTTGAAGTAACCGTCCACAATGACTGCTCTTGTGC 934

    117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
        |||||||||||||||||||:::|||||||||||||||||||||:::|||:
    935 AAAGATGCAAGATCTGCTTAGAAAAGCCTCTGAGGAGGACCACAGCAACT 984

    134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
        ::|||||||||||||||:::|||||||||||||||||||||||:::::|||
    985 CGGCCTGCTTCGCCTGCGTCCTGCTGAGCCACGGGGAAGAGGACCTGATT 1034

    151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1035 TACGGGAAAGATGGCGTGACACCCATAAAGGATCTGACAGCTCATTTTAG 1084

    167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1085 GGGAGACCGATGCAAAACCCTGTTAGAGAAACCCAAACTCTTCTTCATTC 1134

    184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
        ||||||||||||||||||||||||||||:::|||||||||||||||||||
   1135 AGGCATGCCGAGGGACGGAGCTCGACGATGGAATCCAGGCTGACTCGGGG 1184

    201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
        ||||||||||||   ||||||||||||||   ||||||||||||||||||
   1185 CCCATCAACGACATTGACGCTAATCCCCGCAACAAGATCCCGGTGGAAGC 1234

    217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||:
   1235 CGACTTCCTCTTTGCTTACTCCACGGTTCCAGGTTATTACTCATGGAGGA 1284

    234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
        ::||||||:::||||||||||||||||||||||||||||||||||||:::
   1285 ACCCAGGGAAAGGCTCCTGGTTTGTGCAGGCCCTCTGCTCCATCCTGAAT 1334

    251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
        ||||||||||||:::||||||||||||||||||||||||||||||||||
   1335 GAGCATGGCAAGGACCTCGAGATCATGCAGATCCTGACCAGGGTGAACGA 1384

    267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
        ||||||||||||||||||||||||||||||||||||||:::||:::|
   1385 CAGGGTGGCCAGGCACTTCGAGTCCCAGTCTGATGATCCACGCTTCAACG 1434

    284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
        ||||||||||||||||||:::|||||||||||||||||||||||||||||
   1435 AGAAGAAGCAGATCCCGTGTATGGTGTCCATGCTCACCAAAGAGCTGTAC 1484

    301 PheSerGln 303
        ||||||:::
   1485 TTCAGCCGT 1493

seq_name: gb_ro:BC005428

seq_documentation_block:
LOCUS       BC005428               2305 bp    mRNA    linear   ROD 12-JUL-2001

DEFINITION  Mus musculus, caspase 7, clone MGC:5872 IMAGE:3600155, mRNA,
            complete cds.
ACCESSION   BC005428
VERSION     BC005428.1  GI:13529367
KEYWORDS    MGC.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2305)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 11 Row: b Column: 10
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 2094813.
FEATURES             Location/Qualifiers
     source          1. .2305
                     /organism="Mus musculus"
                     /db_xref="LocusID:12369"
                     /db_xref="taxon:10090"
                     /clone="MGC:5872 IMAGE:3600155"
                     /tissue_type="Mammary tumor. Metallothionien-TGF alpha
                     model. 10 month old virgin mouse. Taken by biopsy."
                     /clone_lib="NCI_CGAP_Mam1"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             47. .958
                     /codon_start=1
                     /product="caspase 7"
                     /protein_id="AAH05428.1"
                     /db_xref="GI:13529368"
                     /translation="MTDDQDCAAELEKVDSSSEDGVDAKPDRSSIISSILLKKKRNAS
                     AGPVRTGRDRVPTYLYRMDFQKMGKCIIINNKNFDKATGMDVRNGTDKDAGALFKCFQ
                     NLGFEVTVHNDCSCAKMQDLLRKASEEDHSNSACFACVLLSHGEEDLIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDIDANPRNKIPVEAD
                     FLFAYSTVPGYYSWRNPGKGSWFVQALCSILNEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPRFNEKKQIPCMVSMLTKELYFSR"
BASE COUNT      665 a    523 c    538 g    579 t
ORIGIN


alignment_scores:
            Quality: 1346.00            Length:    303
              Ratio:  4.739               Gaps:      0
 Percent Similarity:  93.729   Percent Identity: 81.518

alignment_block:
 US-09-613-508-2 x BC005428   ..

 Align seg 1/1  to: BC005428  from: 1  to: 2305

      1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
        |||:::|||||||||   |||      |||            ||||||::
     47 ATGACCGATGATCAGGACTGTGCTGCGGAGCTGGAAAAGGTGGATTCTTC 96
```

```
 17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
    ::::||||||:::|||||||||||||||||||||||||||||||:::::   |
 97 CAGCGAAGACGGAGTTGACGCCAAGCCAGACCGCTCCTCTATCATCTCCT 146

 34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
    ||:::::   |||||||||:::|||:::::         :::::|||
147 CTATTCTCTTGAAGAAGAAGAGAAATGCCTCTGCGGGCCCCGTCAGGACC 196

 51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
       ||||||||||||||||||||   |||:::|||:::|||:::|||::
197 GGCCGGGACCGAGTGCCCACTTATCTGTACCGCATGGATTTCCAGAAGAT 246

 67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
    :|||||||||||||||||||||||||||||||||||||:::|||||||
247 GGGTAAATGCATCATCATAAACAACAAGAACTTCGACAAAGCGACAGGTA 296

 84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
    ||   |||||||||||||||||||||||||   ||||||||||||||||
297 TGGACGTCCGGAATGGGACGGACAAAGATGCAGGGGCCCTCTTCAAGTGC 346

101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
    |||::::::|||||||||:::|||   |||::::|||||||||||||||
347 TTCCAAAACCTGGGTTTTGAAGTAACCGTCCACAATGACTGCTCTTGTGC 396

117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
    |||||||||||||||||||:::|||||||||||||||||||||:::|||:
397 AAAGATGCAAGATCTGCTTAGAAAAGCCTCTGAGGAGGACCACAGCAACT 446

134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
    ::|||||||||||||||::|||||||||||||||||||||||:::::|||
447 CGGCCTGCTTCGCCTGCGTCCTGCTGAGCCACGGGGAAGAGGACCTGATT 496

151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
    ||||||||||||||||||||||||||||||||||||||||||||||||||
497 TACGGGAAAGATGGCGTGACACCCATAAAGGATCTGACAGCTCATTTTAG 546

167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
    ||||||||||||||||||||||||||||||||||||||||||||||||||
547 GGGAGACCGATGCAAAACCCTGTTAGAGAAACCCAAACTCTTCTTCATTC 596

184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
    |||||||||||||||||||||||||||||:::|||||||||||||||||
597 AGGCATGCCGAGGGACGGAGCTCGATGATGGAATCCAGGCTGACTCGGGG 646

201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
    ||||||||||||   ||||||||||||||   ||||||||||||||||||
647 CCCATCAACGACATTGACGCTAATCCCCGCAACAAGATCCCGGTGGAAGC 696

217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
    ||||||||||||||||||||||||||||||||||||||||||||||||:
697 CGACTTCCTCTTTGCTTACTCCACGGTTCCAGGTTATTACTCATGGAGGA 746

234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
    ::|||||:::|||||||||||||||||||||||||||||||||||||:::
747 ACCCAGGGAAAGGCTCCTGGTTTGTGCAGGCCCTCTGCTCCATCCTGAAT 796

251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
    ||||||||||||:::|||||||||||||||||||||||||||||||||||
797 GAGCATGGCAAGGACCTCGAGATCATGCAGATCCTGACCAGGGTGAACGA 846

267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
    ||||||||||||||||||||||||||||||||||||||:::|||:::|
847 CAGGGTGGCCAGGCACTTCGAGTCCCAGTCTGATGATCCACGCTTCAACG 896

284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
    ||||||||||||||||||:::|||||||||||||||||||||||||||||
897 AGAAGAAGCAGATCCCGTGTATGGTGTCCATGCTCACCAAAGAGCTGTAC 946

301 PheSerGln 303
    ||||||:::
947 TTCAGCCGT 955
```

```
seq_name: gb_ro:MMCASP7

seq_documentation_block:
LOCUS       MMCASP7                 2372 bp    mRNA    linear   ROD 27-MAY-1997
DEFINITION  M.musculus mRNA for caspase-7.
ACCESSION   Y13088
VERSION     Y13088.1  GI:2094813
KEYWORDS    caspase-7; cysteine protease.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2372)
  AUTHORS   Van de Craen,M., Vandenabeele,P., Declercq,W., Van den Brande,I.,
            Van Loo,G., Molemans,F., Schotte,P., Van Criekinge,W., Beyaert,R.
            and Fiers,W.
  TITLE     Characterization of seven murine caspase family members
  JOURNAL   FEBS Lett. 403 (1), 61-69 (1997)
  MEDLINE   97190206
REFERENCE   2  (bases 1 to 2372)
  AUTHORS   Van de Craen,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAY-1997) M. Van de Craen, University of Gent,
            Laboratory of Molecular Biology, K.L. Ledeganckstraat 35, B-9000
            Gent, BELGIUM
FEATURES             Location/Qualifiers
     source          1. .2372
                     /organism="Mus musculus"
                     /strain="C3H/An"
                     /db_xref="taxon:10090"
                     /cell_line="L929r2"
                     /cell_type="fibrosarcoma"
     CDS             190. .1101
                     /standard_name="casp-7"
                     /function="cysteine protease"
                     /codon_start=1
                     /product="caspase-7"
                     /protein_id="CAA73530.1"
                     /db_xref="GI:2094814"
                     /db_xref="SWISS-PROT:P97864"
                     /translation="MTDDQDCAAELEKVDSSSEDGVDAKPDRSSIISSILLKKKRNAS
                     AGPVRTGRDRVPTYLYRMDFQKMGKCIIINNKNFDKATGMDVRNGTDKDAGALFKCFQ
                     NLGFEVTVHNDCSCAKMQDLLRKASEEDHSNSACFACVLLSHGEEDLIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDIDANPRNKIPVEAD
                     FLFAYSTVPGYYSWRNPGKGSWFVQALCSILNEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPRFNEKKQIPCMVSMLTKELYFSR"
BASE COUNT      608 a    588 c    574 g    602 t
ORIGIN


alignment_scores:
           Quality: 1346.00            Length:    303
             Ratio:  4.739              Gaps:      0
 Percent Similarity:  93.729   Percent Identity: 81.518

alignment_block:
 US-09-613-508-2 x MMCASP7   ..

 Align seg 1/1  to: MMCASP7  from: 1  to: 2372

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         |||:::|||||||||   |||      |||           ||||||::
     190 ATGACCGATGATCAGGACTGTGCTGCGGAGCTGGAAAAGGTGGATTCTTC 239

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ::::||||||:::|||||||||||||||||||||||||||||||:::::   |
     240 CAGCGAAGACGGAGTTGACGCCAAGCCAGACCGCTCCTCTATCATCTCCT 289

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||:::::   |||||||||:::|||:::::         :::::|||
     290 CTATTCTCTTGAAGAAGAAGAGAAATGCCTCTGCGGGCCCCGTCAGGACC 339

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
```

```
      ||||||||||||||||||||   |||:::|||:::|||:::|||::
  340 GGCCGGGACCGAGTGCCCACTTATCTGTACCGCATGGATTTCCAGAAGAT 389

   67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
      :|||||||||||||||||||||||||||||||||||||||:::|||||||
  390 GGGTAAATGCATCATCATAAACAACAAGAACTTCGACAAAGCGACAGGTA 439

   84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
      ||   ||||||||||||||||||||||||||   ||||||||||||||||
  440 TGGACGTCCGGAATGGGACGGACAAAGATGCAGGGGCCCTCTTCAAGTGC 489

  101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
      |||:::::|||||||||:::|||   |||:::||||||||||||||||||
  490 TTCCAAAACCTGGGTTTTGAAGTAACCGTCCACAATGACTGCTCTTGTGC 539

  117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
      |||||||||||||||||:::|||||||||||||||||||||:::|||:
  540 AAAGATGCAAGATCTGCTTAGAAAAGCCTCTGAGGAGGACCACAGCAACT 589

  134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
      ::|||||||||||||||:::|||||||||||||||||||||:::::|||
  590 CGGCCTGCTTCGCCTGCGTCCTGCTGAGCCACGGGGAAGAGGACCTGATT 639

  151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  640 TACGGGAAAGATGGCGTGACACCCATAAAGGATCTGACAGCTCATTTTAG 689

  167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  690 GGGAGACCGATGCAAAACCCTGTTAGAGAAACCCAAACTCTTCTTCATTC 739

  184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
      |||||||||||||||||||||||||||:::|||||||||||||||||||
  740 AGGCATGCCGAGGGACGGAGCTCGATGATGGAATCCAGGCTGACTCGGGG 789

  201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
      ||||||||||||   ||||||||||||||   ||||||||||||||||||
  790 CCCATCAACGACATTGACGCTAATCCCCGCAACAAGATCCCGGTGGAAGC 839

  217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
      ||||||||||||||||||||||||||||||||||||||||||||||||:
  840 CGACTTCCTCTTTGCTTACTCCACGGTTCCAGGTTATTACTCATGGAGGA 889

  234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
      ::||||||:::|||||||||||||||||||||||||||||||||||:::
  890 ACCCAGGGAAAGGCTCCTGGTTTGTGCAGGCCCTCTGCTCCATCCTGAAT 939

  251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
      ||||||||||||:::|||||||||||||||||||||||||||||||||||
  940 GAGCATGGCAAGGACCTCGAGATCATGCAGATCCTGACCAGGGTGAACGA 989

  267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
      ||||||||||||||||||||||||||||||||||||||:::|||:::|
 990 CAGGGTGGCCAGGCACTTCGAGTCCCAGTCTGATGATCCACGCTTCAACG 1039

  284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
      ||||||||||||||||||||:::|||||||||||||||||||||||||||
1040 AGAAGAAGCAGATCCCGTGTATGGTGTCCATGCTCACCAAAGAGCTGTAC 1089

  301 PheSerGln 303
      ||||||:::
1090 TTCAGCCGT 1098
```

```
OM of: US-09-613-508-2 to: N_Geneseq_032802:*   out_format : pfs

Date: Aug 15, 2002  8:52 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO_spool/US09613508/runat_14082002_112648_7265/app_query.fasta_1.701
 -DB=N_Geneseq_032802 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US09613508_@CGN1_1_190 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-613-508-2
Query length: 303
Database: N_Geneseq_032802:*
Database sequences: 1736436
Database length: 858457221
Search time (sec): 267.890000


score_list:
Sequence              Strd Orig    ZScore    EScore Len  ! Documentation  ..
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81223 + 1606.00 3403.09  3.3e-18
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT66992 + 1606.00 3392.69  1.3e-18
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03914 + 1606.00 3392.69  1.3e-1
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT66970 + 1606.00 3392.53  1.3e-18
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT15276 + 1580.00 3343.01  7.4e-17
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT66993 + 1381.50 2912.91  6.7e-15
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13078 +  727.50 1505.02  1.8e-7
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13079 +  727.50 1504.57  1.9e-7
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13080 +  727.50 1504.17  2.0e-7
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13081 +  727.50 1503.78  2.1e-7
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAC64043 +  722.00 1503.71  2.1e-75
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13075 +  721.00 1514.06  5.5e-7
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT33567 +  721.00 1510.37  8.8e-76
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV05471 +  721.00 1510.37  8.8e-76
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT15277 +  721.00 1510.36  8.8e-76
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAI66510 +  720.50 1512.99  6.3e-7
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81219 +  718.00 1507.63  1.3e-75
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAS10503 +  718.00 1494.75  6.5e-7
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03910 +  718.00 1494.75  6.5e-7
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13084 +  717.50 1482.33  3.2e-7
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV32608 +  711.00 1490.65  1.1e-74
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV32615 +  711.00 1490.65  1.1e-74
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV13959 +  709.00 1484.56  2.4e-74
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV13960 +  709.00 1484.56  2.4e-74
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03918 +  707.00 1484.82  2.3e-7
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAS10510 +  703.00 1470.66  1.4e-7
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAH26878 +  556.50 1146.30  1.7e-5
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAV80061 +  552.00 1140.13  3.7e-55
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL03125 +  539.00 1120.69  4.5e-5
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL03124 +  539.00 1110.06  1.8e-5
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL09303 +  532.50 1097.10  9.2e-5
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL09302 +  512.50 1047.63  5.3e-5
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL06952 +  512.50 1043.77  8.6e-5
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81222 +  474.00  985.91  1.4e-46
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAZ98834 +  468.00  969.82  1.1e-45
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAZ98839 +  468.00  969.62  1.2e-45
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT43537 +  466.00  962.55  2.9e-45
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03913 +  466.00  962.55  2.9e-4
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV26613 +  450.00  924.66  3.7e-43
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD12324 +  447.50  923.45  4.3e-4
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:ABL26207 +  446.00  925.57  3.3
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81226 +  444.50  917.42  9.4e
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT90097 +  444.50  915.56  1.2e
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03917 +  444.50  907.35  3.4
/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV53209 +  433.50  894.82  1.7e

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:AAX81223

seq_documentation_block:
ID   AAX81223 standard; DNA; 913 BP.
XX
AC   AAX81223;
XX
DT   10-SEP-1999  (first entry)
XX
DE   Nucleotide sequence of caspase-7 (Mch3).
XX
KW   Rev-caspase; cysteine protease; zymogen; caspase; cancer; gene therapy;
KW   autoimmune disease; caspase-mediated apoptosis; neurodegenerative;
KW   tumour cell; myocardial infarction; human; ds.
XX
OS   Homo sapiens.
XX
PN   WO9935277-A2.
XX
PD   15-JUL-1999.
XX
PF   11-JAN-1999;   99WO-US00632.
XX
PR   09-JAN-1998;   98US-0070987.
XX
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES;
XX
DR   WPI; 1999-419353/35.
DR   P-PSDB; AAY21721.
XX
PT   New isolated nucleic acid molecule encoding a rev-caspase  - used
PT   for screening and identifying inhibitors or enhancers for treating
PT   cancer or autoimmune disease
XX
PS   Disclosure; Fig 17A-B; 74pp; English.
XX
CC   The invention relates to nucleic acid molecules encoding rev-caspases.
CC   Rev-caspases are cysteine proteases that specifically cleave proteins
CC   after Asp residues and is expressed as a zymogen, in which the small
CC   subunit is N-terminal to a large subunit. A gene delivery vehicle
CC   comprising a rev-caspase coding sequence is useful for the treatment of
CC   cancer, where the gene delivery vehicle is internalised by tumour cells.
CC   The gene delivery vehicle can also be used to treat autoimmune diseases.
CC   Cells transfected with a rev-caspase expressing vector can be used in
CC   identification of inhibitors or enhancers of caspase-mediated apoptosis.
CC   In vitro translated rev-caspase can be used to identify an inhibitor or
CC   enhancer of caspase processing activity. Caspase inhibitors are useful
CC   for treating neurodegenerative diseases as well as for inhibiting
CC   apoptosis in the heart following myocardial infarction. Sequences
CC   AAX81217 -AAX81226 represent human caspase genes encoding caspase 1-10
CC   gene products (AAY21715-Y21724).
XX
SQ   Sequence 913 BP; 263 A; 218 C; 224 G; 207 T; 1 other;


alignment_scores:
           Quality: 1606.00              Length:    303
             Ratio:  5.300                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.340

alignment_block:
 US-09-613-508-2 x AAX81223   ..

 Align seg 1/1  to: AAX81223  from: 1  to: 913
```

```
  1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
    |||||||||||||||||||||||||||||||||||||||||||||||||||
  1 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 50

 17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
    |||||||||||||||||||||||||||||||||||||||||||||||||||
 51 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 100

 34 erLeuPheSerLysLysLysAsnValThrMetArgSerIleLysThr 50
    |||||||||||||||||||||||||||||||||||||||||||||||||||
101 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 150

 51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
    |||||||||||||||||||||||||||||||||||||||||||||||||||
151 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 200

 67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
    |||||||||||||||||||||||||||||||||||||||||||||||||||
201 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 250

 84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
    |||||||||||||||||||||||||||||||||||||||||||||||||||
251 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 300

101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
    |||||||||||||||||||||||||||||||||||||||||||||||||||
301 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 350

117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
    |||||||||||||||||||||||||||||||||||||||||||||||||||
351 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 400

134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
    |||||||||||||||||||||||||||||||||||||||||||||||||||
401 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 450

151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
    |||||||||||||||||||||||||||||||||||||||||||||||||||
451 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 500

167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
    |||||||||||||||||||||||||||||||||||||||||||||||||||
501 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 550

184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
    |||||||||||||||||||||||||||:::|||||||||||||||||||
551 AGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGG 600

201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
    |||||||||||||||||||||||||||||||||||||||||||||||||||
601 CCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGC 650

217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
    |||||||||||||||||||||||||||||||||||||||||||||||||||
651 TGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGA 700

234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
    |||||||||||||||||||||||||||||||||||||||||||||||||||
701 GCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAG 750

251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
    ||||||||||:::|||||||||||||||||||||||||||||||||||||
751 GAGCACGGAAAAGACCTGGAAATCATGCAGATCCTCACCAGGGTGAATGA 800

267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
    |||||||||||||||||||||||||||||||||||||||||||||||||||
801 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 850

284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
    |||||||||||||||||||||||||||||||||||||||||||||||||||
851 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTAC 900

301 PheSerGln 303
```

```
    |||||||||
901 TTCAGTCAA 909
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT66992

seq_documentation_block:

```
ID   AAT66992 standard; cDNA; 2309 BP.
XX
AC   AAT66992;
XX
DT   05-AUG-1997  (first entry)
XX
DE   Apototic protease Mch3-alpha cDNA.
XX
KW   Mch3-alpha; cysteine protease; apoptosis; AIDS; ischaemia;
KW   neurodegenerative disease; gene therapy; diagnosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             44..955
FT                   /*tag= a
FT   intron          595..596
FT                   /*tag= b
FT                   /note= "site of intron in genomic DNA"
XX
PN   WO9718313-A1.
XX
PD   22-MAY-1997.
XX
PF   12-NOV-1996;   96WO-US18118.
XX
PR   13-NOV-1995;   95US-0556627.
XX
PA   (IDUN-) IDUN PHARM INC.
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES,  Armstrong R,  Fernandes-Alnemri T,  Litwack G;
PI   Tomaselli K;
XX
DR   WPI; 1997-289289/26.
DR   P-PSDB; AAW15262.
XX
PT   New gene encoding Mch3, a cysteine protease that regulates apoptosis
PT   - for treating human diseases associated with apoptosis, and
PT   screening for antagonists and agonists of Mch3
XX
PS   Claim 2; Fig 1; 52pp; English.
XX
CC   A cDNA clone (AAT66992) codes for Mch3-alpha (AAW15262), an apoptotic
CC   protein and novel member of the interleukin-1-beta converting
CC   enzyme (ICE) family of cysteine proteases.  It was obtd. by
CC   screening a human database of expressed sequence tags for sequences
CC   showing homology to ICE family members, and use of a partial cDNA
CC   clone with high homology to CPP32 and Ced-3 to screen a Jurkat cDNA
CC   library.  A second cDNA clone, Mch3-beta (AAT66993), contains a
CC   deletion and insertion corresponding to nucleotides 488-592 of the
CC   Mch3-alpha sequence.  Mch3 nucleic acids can be used to produce
CC   recombinant Mch3 polypeptides useful for treating diseases
CC   associated with cell death, such as AIDS, ischaemic injury,
CC   neurodegenerative diseases, etc.  They can also be used in gene
CC   therapy to regulate apoptosis.
XX
SQ   Sequence 2309 BP; 680 A; 453 C; 531 G; 645 T; 0 other;
```

alignment_scores:

```
           Quality: 1606.00              Length:    303
             Ratio:  5.300                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.340
```

alignment_block:

```
 US-09-613-508-2 x AAT66992    ..
```

```
Align seq 1/1  to: AAT66992  from: 1  to: 2309

      1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     44 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 93

     17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     94 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 143

     34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    144 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 193

     51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    194 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 243

     67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    244 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 293

     84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    294 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 343

    101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    344 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 393

    117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    394 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 443

    134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    444 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 493

    151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    494 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 543

    167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    544 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 593

    184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
        ||||||||||||||||||||||||||||||:::||||||||||||||||||
    594 AGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGG 643

    201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    644 CCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGC 693

    217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    694 TGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGA 743

    234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    744 GCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAG 793

    251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
        ||||||||||||:::||||||||||||||||||||||||||||||||||
    794 GAGCACGGAAAAGACCTGGAAATCATGCAGATCCTCACCAGGGTGAATGA 843

    267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    844 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 893

    284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||

    894 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTAC 943

    301 PheSerGln 303
        ||||||||
    944 TTCAGTCAA 952

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD03914

seq_documentation_block:
ID   AAD03914 standard; DNA; 2309 BP.
XX
AC   AAD03914;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Human caspase-7 DNA.
XX
KW   Human; caspase-7; interleukin-1 converting enzyme; ICE-LAP3; CHM-1; Mhc3;
KW   cysteine protease; apoptosis; caspase expression cassette; metastasis;
KW   tumour; cathepsin B; urokinase; proliferation; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             44..955
FT                   /*tag= a
FT                   /product= "Human caspase-7"
XX
PN   WO200129232-A2.
XX
PD   26-APR-2001.
XX
PF   19-OCT-2000; 2000WO-US28941.
XX
PR   20-OCT-1999;   99US-0160559.
PR   14-AUG-2000; 2000US-0225564.
XX
PA   (SCIO-) SCIOS INC.
XX
PI   Cordell B,  Li Y;
XX
DR   WPI; 2001-290920/30.
DR   P-PSDB; AAE00604.
XX
PT   Novel fusion polypeptide comprising first and second caspase subunit
PT   separated by cleavage site not associated in nature with caspase
PT   subunit, useful for cloning gene encoding enzymes involved in
PT   proteolytic cleavage
XX
PS   Disclosure; Fig 13; 116pp; English.
XX
CC   The present sequence is a DNA encoding human Caspase-7 also known as
CC   interleukin-1 converting enzyme (ICE) LAP3, CHM-1 and Mhc3. Caspases are
CC   a family of cysteine proteases, that participate in the initiation and
CC   execution of apoptosis. Caspases exist as pro-enzymes, activated by
CC   cleavage into a large and small subunit, occurring after specific
CC   aspartic acid residues within the pro-enzyme sequence.
CC   The present invention relates to a method for functional cloning of genes
CC   encoding proteins or enzymes involved in proteolytic cleavage. The
CC   invention is based on the use of caspase expression cassettes comprising
CC   the coding sequence of a proteolytic cleavage site flanked by sequences
CC   encoding two caspase subunits. A fusion polypeptide comprising a first
CC   and a second caspase subunit, separated by a cleavage site not associated
CC   in nature, is useful for cloning gene encoding enzymes involved in
CC   proteolytic cleavage. An expression cassette containing fusion
CC   polypeptide is used to identify a mutant cell line deficient in an
CC   enzyme of interest and is also useful for diagnosis and suppression of
CC   proliferation or metastases of a tumour cell characterised by
CC   overexpression of a polypeptide (e.g. Cathepsin B or urokinase,
CC   selectively expressed in the tumour cells). DNA encoding fusion
CC   polypeptide is used in gene therapy.
XX
SQ   Sequence 2309 BP; 680 A; 453 C; 531 G; 645 T; 0 other;
```

```
alignment_scores:
            Quality: 1606.00              Length:    303
              Ratio:  5.300                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.340

alignment_block:
 US-09-613-508-2 x AAD03914   ..

 Align seg 1/1  to: AAD03914  from: 1  to: 2309

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      44 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 93

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      94 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 143

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     144 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 193

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     194 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 243

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     244 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 293

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     294 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 343

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     344 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 393

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     394 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 443

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     444 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 493

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     494 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 543

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     544 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 593

     184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
         ||||||||||||||||||||||||||||||||:::||||||||||||||||
     594 AGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGG 643

     201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     644 CCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGC 693

     217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     694 TGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGA 743

     234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     744 GCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAG 793

     251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
         ||||||||||||:::||||||||||||||||||||||||||||||||||||
     794 GAGCACGGAAAAGACCTGGAAATCATGCAGATCCTCACCAGGGTGAATGA 843

     267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     844 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 893

     284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     894 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTAC 943

     301 PheSerGln 303
         ||||||||
     944 TTCAGTCAA 952

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT66970

seq_documentation_block:
ID   AAT66970 standard; cDNA; 2341 BP.
XX
AC   AAT66970;
XX
DT   21-JUL-1997  (first entry)
XX
DE   Cysteine protease CMH-1 cDNA.
XX
KW   Cysteine protease; CMH-1; CPP32/Mch2 homologue-1; apoptosis;
KW   programmed cell death; cancer; neurodegenerative disease;
KW   autoimmune disease; gene therapy; diagnosis; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             54..965
FT                   /*tag= a
FT                   /product= CMH-1 proenzyme
FT                   /note= "bases 123-962 and 138-962 encode active
FT                           active CMH-1 (Claim 2)"
XX
PN   WO9716552-A1.
XX
PD   09-MAY-1997.
XX
PF   01-NOV-1996;   96WO-US17431.
XX
PR   16-NOV-1995;   95US-0558733.
PR   03-NOV-1995;   95US-0007211.
PR   06-NOV-1995;   95US-0007251.
XX
PA   (VERT-) VERTEX PHARM INC.
XX
PI   Lippke JA,  Su M;
XX
DR   WPI; 1997-272121/24.
DR   P-PSDB; AAW15247.
XX
PT   DNA encoding active, activatable or inactive cysteine protease CMH-1
PT   - useful in gene therapy for promoting and inhibiting apoptosis and
PT   for diagnosing cells with potential for apoptosis
XX
PS   Claim 1; Page 30-32; 48pp; English.
XX
CC   A cDNA clone (T66970) includes to coding sequence for the proenzyme
CC   (W15247) of CMH-1, a human cysteine protease involved in apoptosis.
CC   To obtain the clone, the GenBank Expressed Sequence Tag (EST)
CC   database was screened for clones encoding polypeptides similar to
CC   interleukin-1b converting enzyme.  Partial cDNA clone T50828 was
CC   identified.  Partial cDNA clone 72778 corresponding to this EST was
CC   then obtd.  A spleen cDNA pool was subjected to PCR using primers
CC   (T66971-92) based on a 5' adaptor and EST T50828.  A PCR fragment
CC   contg. the 5' coding region was ligated to the clone 72778 3' region
CC   to obtain the full-length CHM-1 sequence.  CHM-1 (i.e. CPP32/Mch2
CC   homologue-1) DNA can be used in gene therapy to promote apoptosis,
```

```
CC   esp. in cancer cells, and to modulate patient cholesterol levels,
CC   as well as to diagnose cells with potential for apoptosis.
CC   Antisense sequences or sequences encoding inactive CHM-1 mutants
CC   can be used to inhibit apoptosis, e.g. to prevent neurological
CC   disease.
XX
SQ   Sequence 2341 BP; 702 A; 455 C; 536 G; 648 T; 0 other;


alignment_scores:
           Quality: 1606.00              Length:    303
             Ratio:  5.300                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.340

alignment_block:
 US-09-613-508-2 x AAT66970   ..

 Align seg 1/1  to: AAT66970  from: 1  to: 2341

      1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     54 ATGGCAGATGATCAGGGCTGTATTGAGGAGCAGGGGGTTGAGGATTCAGC 103

     17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    104 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 153

     34 erLeuPheSerLysLysLysAsnValThrMetArgSerIleLysThr 50
        ||||||||||||||||||||||||||||||||||||||||||||||||
    154 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 203

     51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    204 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 253

     67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    254 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 303

     84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    304 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 353

    101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    354 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 403

    117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    404 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 453

    134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    454 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 503

    151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    504 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 553

    167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    554 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 603

    184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
        ||||||||||||||||||||||||||||||:::|||||||||||||||||
    604 AGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGG 653

    201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    654 CCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGC 703

    217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||

    704 TGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGA 753

    234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    754 GCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAG 803

    251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
        ||||||||||||:::|||||||||||||||||||||||||||||||||||
    804 GAGCACGGAAAAGACCTGGAAATCATGCAGATCCTCACCAGGGTGAATGA 853

    267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    854 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 903

    284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    904 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTAC 953

    301 PheSerGln 303
        ||||||||
    954 TTCAGTCAA 962

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT15276

seq_documentation_block:
ID   AAT15276 standard; cDNA; 1371 BP.
XX
AC   AAT15276;
XX
DT   27-OCT-1996  (first entry)
XX
DE   Human interleukin-I-converting enzyme-like apoptosis protease-3 cDNA.
XX
KW   ICE-LAP-3; interleukin-I-converting enzyme-like apoptosis protease;
KW   enzyme; Alzheimer's disease; Parkinson's disease; septic shock;
KW   head injury; rheumatoid arthritis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             115..1137
FT                   /*tag= a
XX
PN   WO9613603-A1.
XX
PD   09-MAY-1996.
XX
PF   06-JUN-1995;   95WO-US07235.
XX
PR   01-NOV-1994;   94US-0334251.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Hastings GA,  He W,  Hudson PL,  Rosen CA;
XX
DR   WPI; 1996-239509/24.
DR   P-PSDB; AAR95830.
XX
PT   Human ICE-LAP-3 and -4 DNA and protein  - useful in the diagnosis
PT   and treatment of Alzheimer's disease, Parkinson's disease,
PT   rheumatoid arthritis, septic shock and head injury
XX
PS   Claim 1; Page 91-92; 67pp; English.
XX
CC   This sequence was derived from a cDNA library from human endometrial
CC   tumuor and may be used for the recombinant production of the ICE-LAP-
CC   3 protein e.g. in COS cells.  The encoded protein may be used
CC   therapeutically, e.g. as an antitumor or antiviral agent and to
CC   control embryonic development and tissue homeostasis.  The protein
CC   can also be used to treat immunosuppression disorders, such as AIDS,
CC   by targeting virus infected cells for cell death.  The DNA may find
CC   use in gene therapy applications.
XX
```

```
SQ   Sequence 1371 BP; 355 A; 357 C; 354 G; 305 T; 0 other;

alignment_scores:
           Quality: 1580.00              Length:    305
             Ratio:  5.215                Gaps:      2
 Percent Similarity:  99.344   Percent Identity: 98.689

alignment_block:
 US-09-613-508-2 x AAT15276   ..

 Align seg 1/1  to: AAT15276  from: 1  to: 1371

       1 MetAlaAspAsp.GlnGlyCysIleGluGluGlnGlyValGluAspSerA 17
         |||||||||||| ||||||||||||||||||||||||||||||||||||||
     339 ATGGCAGATGATTCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAG 388

      17 laAsnGluAspSerValAspAlaLysProAspArgSerSerPheValPro 33
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     389 CAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCG 438

      34 SerLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysTh 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     439 TCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGAC 488

      50 rThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysL 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     489 CACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGC 538

      67 euGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGly 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     539 TGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGT 588

      84 MetGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCy 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     589 ATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTG 638

     100 sPheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysA 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     639 CTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTG 688

     117 laLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsn 133
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     689 CCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAAT 738

     134 AlaAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIl 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     739 GCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAAT 788

     150 eTyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheA 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     789 TTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTA 838

     167 rgGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIle 183
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     839 GGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATT 888

     184 GlnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGl 200
         ||||||||||||||||||||||||||||||:::|||||||||||||||||
     889 CAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGG 938

     200 yProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluA 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     939 GCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAG 988

     217 laAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArg 233
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     989 CTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGG 1038

     234 SerProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGl 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1039 AGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGA 1088

     250 uGluHisGlyLysGluLeuGluIleMetGlnIleLeu.ThrArgValAsn 266
         ||||||||||||||:::|||||||||||||||||||| ||||||||||||
    1089 GGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAAT 1138

     267 AspArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHi 283
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1139 GACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCA 1188

     283 sGluLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuT 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1189 TGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCT 1238

     300 yrPheSerGln 303
         ||||||||||
    1239 ACTTCAGTCAA 1249

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:AAT66993

seq_documentation_block:
ID   AAT66993 standard; cDNA; 2377 BP.
XX
AC   AAT66993;
XX
DT   05-AUG-1997  (first entry)
XX
DE   Cysteine protease Mch3-beta cDNA.
XX
KW   Mch3-beta; cysteine protease; apoptosis; AIDS; ischaemia;
KW   neurodegenerative disease; gene therapy; diagnosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             146..839
FT                   /*tag= a
XX
PN   WO9718313-A1.
XX
PD   22-MAY-1997.
XX
PF   12-NOV-1996;   96WO-US18118.
XX
PR   13-NOV-1995;   95US-0556627.
XX
PA   (IDUN-) IDUN PHARM INC.
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES,  Armstrong R,  Fernandes-Alnemri T,  Litwack G;
PI   Tomaselli K;
XX
DR   WPI; 1997-289289/26.
DR   P-PSDB; AAW15263.
XX
PT   New gene encoding Mch3, a cysteine protease that regulates apoptosis
PT   - for treating human diseases associated with apoptosis, and
PT   screening for antagonists and agonists of Mch3
XX
PS   Claim 2; Fig 1; 52pp; English.
XX
CC   A cDNA clone (AAT66993) codes for Mch3-beta (AAW15263), a novel member
CC   of the interleukin-1-beta converting enzyme (ICE) family of
CC   cysteine proteases.  It was obtd. by screening a human database of
CC   expressed sequence tags for sequences showing homology to ICE
CC   family members, and use of a partial cDNA clone with high homology
CC   to CPP32 and Ced-3 to screen a Jurkat cDNA library.  Mch3-beta cDNA
CC   contains a deletion and insertion corresponding to nucleotides
CC   488-592 of Mch3-alpha cDNA (AAT66992) and also has a longer 3'
CC   nontranslated sequence.  Mch3-beta lacks the active site of Mch3-
CC   alpha.  Mch3 nucleic acids can be used to produce recombinant Mch3
CC   polypeptides useful for treating diseases associated with cell
CC   death, such as AIDS, ischaemic injury, neurodegenerative diseases,
CC   etc.  They can also be used in gene therapy to regulate apoptosis.
XX
```

```
SQ   Sequence 2377 BP; 680 A; 485 C; 565 G; 647 T; 0 other;

alignment_scores:
           Quality: 1381.50              Length:    305
             Ratio:  4.934                Gaps:        3
 Percent Similarity:  91.803   Percent Identity: 88.525

alignment_block:
 US-09-613-508-2 x AAT66993   ..

 Align seg 1/1  to: AAT66993  from: 1  to: 2377

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     146 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 195

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     196 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 245

      34 erLeuPheSerLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     246 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 295

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     296 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 345

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     346 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 395

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     396 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 445

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     446 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 495

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     496 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 545

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         ||||||||||||||||||||||||||||||||||||||||||
     546 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAAT..... 589

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
                              :::::::::   ::::::
     590 .....ATGGAATCTTGCTCTGTCACCCAGGC 615

     167 gGlyAspArgCysLysThrLeu......LeuGluLysProLysLeuPheP 182
          |||   :::   :::   |||       |||:::   |||
     616 TGGAGTGCAGCGGCGTGATCTCGGAAGACTGCAACCTCCACCTCC..... 660

     182 heIleGlnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAsp 198
               ||||||||||||||||||||||||||||:::||||||||||||
     661 .....CAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGAC 705

     199 SerGlyProIleAsnAspThrAspAlaAsnProArgTyrLysIleProVa 215
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     706 TCGGGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGT 755

     215 lGluAlaAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerT 232
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     756 GGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGT 805

     232 rpArgSerProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIle 248
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     806 GGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATC 855

     249 LeuGluGluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgVa 265
         ||||||||||||||||||:::||||||||||||||||||||||||||||
     856 CTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCTCACCAGGGT 905

     265 lAsnAspArgValAlaArgHisPheGluSerGlnSerAspAspProHisP 282
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     906 GAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACT 955

     282 heHisGluLysLysGlnIleProCysValValSerMetLeuThrLysGlu 298
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     956 TCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAA 1005

     299 LeuTyrPheSerGln 303
         |||||||||||||||
    1006 CTCTACTTCAGTCAA 1020

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13078

seq_documentation_block:
ID   AAD13078 standard; DNA; 6436 BP.
XX
AC   AAD13078;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-1XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-1XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain  -
XX
PS   Example 4; Page 133-134; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-1XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domain-Mammalia caspase-3 fusion protein.
```

```
XX
SQ   Sequence 6436 BP; 1577 A; 1606 C; 1607 G; 1646 T; 0 other;

alignment_scores:
            Quality:  727.50              Length:    292
              Ratio:  3.277                Gaps:      4
 Percent Similarity:  76.027    Percent Identity: 50.342

alignment_block:
 US-09-613-508-2 x AAD13078   ..

 Align seg 1/1  to: AAD13078  from: 1  to: 6436

      11 GlnGlyValGluAspSerAlaAsn...GluAspSerValAspAlaLysPr 26
         :::|||||||:::   ::::::|||   |||:::|||||||||:::|||
    1123 AGAGGTGTGCACCACGGATCCAACACTGAAAACTCAGTGGATTCAAAATC 1172

      26 oAspArgSerSerPheValProSerLeuPheSerLysLysLysLysAsnV 43
            :::   ::::::   |||:::::::::      :::::::::
    1173 CATTAAA...AATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGG 1219

      43 alThrMetArgSerIleLysThrThrArgAspArgValProThrTyrGln 59
           :::      |||:::   ::::::
    1220 ACTCTGGAATATCCCTGGACAACAGT..................... 1245

      60 TyrAsnMetAsnPheGluLysLeuGlyLysCysIleIleIleAsnAsnLy 76
         |||:::|||::::::   ::::::|||   ||||||||||||||||||
    1246 TATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAA 1295

      76 sAsnPheAspLysValThrGlyMetGlyValArgAsnGlyThrAspLysA 93
         |||||||   |||   |||||||||      |||:::|||||||||   |
    1296 GAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCG 1345

      93 spAlaGluAlaLeuPheLysCysPheArgSerLeuGlyPheAspValIle 109
         |||||      |||   :::   ||||||:::|||   ::::::|||
    1346 ATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 1395

     110 ValTyrAsnAspCysSerCysAlaLysMetGlnAspLeuLeuLysLysAl 126
            ||||||   :::      ::::::   :::|||::::::   ::
    1396 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGT 1445

     126 aSerGluGluAspHisThrAsnAlaAlaCysPheAlaCysIleLeuLeuS 143
         :|||:::|||||||||::::::   :::   |||:::|||:::|||||||
    1446 TTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGA 1495

     143 erHisGlyGluGluAsnValIleTyrGlyLysAspGlyValThrProIle 159
         |||||||||||||||:::::|||:::|||   :::|||   :::   :::
    1496 GCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTG 1545

     160 LysAspLeuThrAlaHisPheArgGlyAspArgCysLysThrLeuLeuGl 176
         |||   :::|||      |||||||||||||||||:::::::|||
    1546 AAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGG 1595

     176 uLysProLysLeuPhePheIleGlnAlaCysArgGlyThrGluLeuAspA 193
         ||||||||||||||:::||||||||||||||||||||||||||||||||
    1596 AAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACT 1645

     193 spAlaIleGlnAlaAspSerGlyProIleAsnAspThrAspAlaAsnPro 209
         :::|||::::::|||||||||      :::|||      :::
    1646 GTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGT..... 1689

     210 ArgTyrLysIleProValGluAlaAspPheLeuPheAlaTyrSerThrVa 226
         :::||||||||||||||||||||||||||||||:::|||||||||||::
    1690 ...CATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGC 1736

     226 lProGlyTyrTyrSerTrpArgSerProGlyArgGlySerTrpPheValG 243
         :|||||||||||||||||||:::         |||||||||||:::|
    1737 ACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCC 1786

     243 lnAlaLeuCysSerIleLeuGluGluHisGlyLysGluLeuGluIleMet 259
         ||:::|||||::::::|||:::::::::::   :::||||||::|||
    1787 AGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATG 1836

     260 GlnIleLeuThrArgValAsnAspArgValAlaArgHisPheGluSerGl 276
         :::|||||||||||||||||||   :::||||||   :::|||||||||
    1837 CACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTT 1886

     276 nSerAspAspProHisPheHisGluLysLysGlnIleProCysValValS 293
          |||   |||      ||||||   ||||||||||||||||||:::||||
    1887 TTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTT 1936

     293 erMetLeuThrLysGluLeuTyrPhe 301
         |||||||||||||||||||||||||
    1937 CCATGCTCACAAAAGAACTCTATTTT 1962

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13079

seq_documentation_block:
ID   AAD13079 standard; DNA; 6703 BP.
XX
AC   AAD13079;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-2XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-2XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain  -
XX
PS   Example 4; Page 134-135; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-2XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domains-Mammalia caspase-3
CC   fusion protein.
XX
SQ   Sequence 6703 BP; 1653 A; 1677 C; 1674 G; 1699 T; 0 other;
```

```
alignment_scores:
            Quality:  727.50              Length:    292
              Ratio:  3.277                Gaps:      4
 Percent Similarity:  76.027   Percent Identity: 50.342

alignment_block:
 US-09-613-508-2 x AAD13079   ..

 Align seg 1/1  to: AAD13079  from: 1  to: 6703

      11 GlnGlyValGluAspSerAlaAsn...GluAspSerValAspAlaLysPr 26
         :::||||||:::   ::::::|||   |||:::|||||||||:::|||
    1390 AGAGGTGTGCACCACGGATCCAACACTGAAAACTCAGTGGATTCAAAATC 1439

      26 oAspArgSerSerPheValProSerLeuPheSerLysLysLysLysAsnV 43
            :::   ::::::   |||:::::::::      :::::::::
    1440 CATTAAA...AATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGG 1486

      43 alThrMetArgSerIleLysThrThrArgAspArgValProThrTyrGln 59
         :::      |||:::   ::::::
    1487 ACTCTGGAATATCCCTGGACAACAGT..................... 1512

      60 TyrAsnMetAsnPheGluLysLeuGlyLysCysIleIleIleAsnAsnLy 76
         |||:::|||::::::   ::::::|||   ||||||||||||||||||||
    1513 TATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAA 1562

      76 sAsnPheAspLysValThrGlyMetGlyValArgAsnGlyThrAspLysA 93
         |||||||   |||   ||||||||      |||:::|||||||||   |
    1563 GAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCG 1612

      93 spAlaGluAlaLeuPheLysCysPheArgSerLeuGlyPheAspValIle 109
         |||||      |||   :::   ||||||:::|||   :::::::|||
    1613 ATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 1662

     110 ValTyrAsnAspCysSerCysAlaLysMetGlnAspLeuLeuLysLysAl 126
            ||||||   :::      ::::::   :::|||::::::  ::
    1663 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGT 1712

     126 aSerGluGluAspHisThrAsnAlaAlaCysPheAlaCysIleLeuLeuS 143
         :|||:::||||||||||:::::   :::   |||:::|||:::|||||||
    1713 TTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGA 1762

     143 erHisGlyGluGluAsnValIleTyrGlyLysAspGlyValThrProIle 159
         ||||||||||||||:::::::|||:::|||   :::|||   :::   :::
    1763 GCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTG 1812

     160 LysAspLeuThrAlaHisPheArgGlyAspArgCysLysThrLeuLeuGl 176
         |||   :::|||      ||||||||||||||||||:::::|||
    1813 AAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGG 1862

     176 uLysProLysLeuPhePheIleGlnAlaCysArgGlyThrGluLeuAspA 193
          |||||||||||||:::|||||||||||||||||||||||||||||||||
    1863 AAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACT 1912

     193 spAlaIleGlnAlaAspSerGlyProIleAsnAspThrAspAlaAsnPro 209
         :::|||::::::|||||||      :::|||      :::
    1913 GTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGT..... 1956

     210 ArgTyrLysIleProValGluAlaAspPheLeuPheAlaTyrSerThrVa 226
         :::|||||||||||||||||||||||||||:::|||||||||||||::
    1957 ...CATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGC 2003

     226 lProGlyTyrTyrSerTrpArgSerProGlyArgGlySerTrpPheValG 243
         :|||||||||||||||||||:::        ||||||||||||:::|
    2004 ACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCC 2053

     243 lnAlaLeuCysSerIleLeuGluGluHisGlyLysGluLeuGluIleMet 259
         ||:::|||||:::::|||:::::::::::   :::||||||:::|||
    2054 AGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATG 2103

     260 GlnIleLeuThrArgValAsnAspArgValAlaArgHisPheGluSerGl 276
         :::||||||||||||||||||   :::|||||||   :::|||||||||
    2104 CACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTT 2153

     276 nSerAspAspProHisPheHisGluLysLysGlnIleProCysValVals 293
          |||   |||      ||||||   |||||||||||||||||:::||||
    2154 TTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTT 2203

     293 erMetLeuThrLysGluLeuTyrPhe 301
         |||||||||||||||||||||||||
    2204 CCATGCTCACAAAAGAACTCTATTTT 2229

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13080

seq_documentation_block:
ID   AAD13080 standard; DNA; 6946 BP.
XX
AC   AAD13080;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-3XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-3XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain  -
XX
PS   Example 4; Page 136-137; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-3XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domains-Mammalia caspase-3
CC   fusion protein.
XX
SQ   Sequence 6946 BP; 1723 A; 1742 C; 1735 G; 1746 T; 0 other;
```

```
alignment_scores:
            Quality:  727.50              Length:    292
              Ratio:  3.277                 Gaps:      4
 Percent Similarity:  76.027    Percent Identity: 50.342

alignment_block:
 US-09-613-508-2 x AAD13080   ..

 Align seg 1/1  to: AAD13080  from: 1  to: 6946

      11 GlnGlyValGluAspSerAlaAsn...GluAspSerValAspAlaLysPr 26
         :::||||||:::   ::::::|||   |||:::|||||||||:::|||
    1633 AGAGGTGTGCACCACGGATCCAACACTGAAAACTCAGTGGATTCAAAATC 1682

      26 oAspArgSerSerPheValProSerLeuPheSerLysLysLysLysAsnV 43
            :::   ::::::   |||::::::::::      :::::::::
    1683 CATTAAA...AATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGG 1729

      43 alThrMetArgSerIleLysThrThrArgAspArgValProThrTyrGln 59
          :::     |||:::   ::::::
    1730 ACTCTGGAATATCCCTGGACAACAGT..................... 1755

      60 TyrAsnMetAsnPheGluLysLeuGlyLysCysIleIleIleAsnAsnLy 76
         |||:::|||::::::   :::::::|||   ||||||||||||||||||
    1756 TATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAA 1805

      76 sAsnPheAspLysValThrGlyMetGlyValArgAsnGlyThrAspLysA 93
         |||||||   |||   |||||||||      |||:::||||||||   |
    1806 GAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCG 1855

      93 spAlaGluAlaLeuPheLysCysPheArgSerLeuGlyPheAspValIle 109
         |||||      |||   :::   |||||||:::|||   :::::::|||
    1856 ATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 1905

     110 ValTyrAsnAspCysSerCysAlaLysMetGlnAspLeuLeuLysLysAl 126
             ||||||   :::       :::::::   :::|||::::::   ::
    1906 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGT 1955

     126 aSerGluGluAspHisThrAsnAlaAlaCysPheAlaCysIleLeuLeuS 143
         :|||:::||||||||||::::::   :::   |||:::|||:::|||||||
    1956 TTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGA 2005

     143 erHisGlyGluGluAsnValIleTyrGlyLysAspGlyValThrProIle 159
         ||||||||||||||:::::::|||:::|||   :::|||   :::   :::
    2006 GCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTG 2055

     160 LysAspLeuThrAlaHisPheArgGlyAspArgCysLysThrLeuLeuGl 176
         |||   :::|||       |||||||||||||||||:::::|||
    2056 AAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGG 2105

     176 uLysProLysLeuPhePheIleGlnAlaCysArgGlyThrGluLeuAspA 193
          ||||||||||||||:::|||||||||||||||||||||||||||||||||
    2106 AAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACT 2155

     193 spAlaIleGlnAlaAspSerGlyProIleAsnAspThrAspAlaAsnPro 209
           :::|||::::::|||||||||      :::|||   :::
    2156 GTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGT...... 2199

     210 ArgTyrLysIleProValGluAlaAspPheLeuPheAlaTyrSerThrVa 226
            :::|||||||||||||||||||||||||||:::||||||||||||::
    2200 ...CATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGC 2246

     226 lProGlyTyrTyrSerTrpArgSerProGlyArgGlySerTrpPheValG 243
         :|||||||||||||||||||:::          ||||||||||||:::|
    2247 ACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCC 2296

     243 lnAlaLeuCysSerIleLeuGluGluHisGlyLysGluLeuGluIleMet 259
         ||:::||||||::::::|||:::::::::::::   :::||||||:::|||
    2297 AGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATG 2346

     260 GlnIleLeuThrArgValAsnAspArgValAlaArgHisPheGluSerGl 276
         :::||||||||||||||||||   :::||||||   :::||||||||||
    2347 CACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTT 2396

     276 nSerAspAspProHisPheHisGluLysLysGlnIleProCysValValS 293
            |||   |||      ||||||   |||||||||||||||||:::||||
    2397 TTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTT 2446

     293 erMetLeuThrLysGluLeuTyrPhe 301
         |||||||||||||||||||||||||
    2447 CCATGCTCACAAAAGAACTCTATTTT 2472

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13081

seq_documentation_block:
ID   AAD13081 standard; DNA; 7189 BP.
XX
AC   AAD13081;
XX
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-4XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ublquitin; cyclic; circular; plasmid pcDNA3-4XUb-C3;
KW   fusion protein; ds.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Homo sapiens.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain -
XX
PS   Example 4; Page 137-138; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-4XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domains-Mammalia caspase-3
CC   fusion protein.
XX
SQ   Sequence 7189 BP; 1793 A; 1807 C; 1796 G; 1793 T; 0 other;


alignment_scores:
            Quality:  727.50              Length:    292
```

```
            Ratio:  3.277              Gaps:      4
 Percent Similarity:  76.027   Percent Identity: 50.342

alignment_block:
 US-09-613-508-2 x AAD13081   ..

 Align seg 1/1  to: AAD13081  from: 1  to: 7189

      11 GlnGlyValGluAspSerAlaAsn...GluAspSerValAspAlaLysPr 26
         :::||||||:::   :::::::|||   |||:::|||||||||:::|||
    1876 AGAGGTGTGCACCACGGATCCAACACTGAAAACTCAGTGGATTCAAAATC 1925

      26 oAspArgSerSerPheValProSerLeuPheSerLysLysLysLysAsnV 43
            :::   ::::::   |||:::::::::      :::::::::
    1926 CATTAAA...AATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGG 1972

      43 alThrMetArgSerIleLysThrThrArgAspArgValProThrTyrGln 59
           :::      |||:::   ::::::
    1973 ACTCTGGAATATCCCTGGACAACAGT..................... 1998

      60 TyrAsnMetAsnPheGluLysLeuGlyLysCysIleIleIleAsnAsnLy 76
         |||:::|||::::::   :::::::|||   ||||||||||||||||||
    1999 TATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAA 2048

      76 sAsnPheAspLysValThrGlyMetGlyValArgAsnGlyThrAspLysA 93
         |||||||   |||   |||||||||      |||:::|||||||||   |
    2049 GAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCG 2098

      93 spAlaGluAlaLeuPheLysCysPheArgSerLeuGlyPheAspValIle 109
         |||||      |||   :::   ||||||:::|||   :::::::|||
    2099 ATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 2148

     110 ValTyrAsnAspCysSerCysAlaLysMetGlnAspLeuLeuLysLysAl 126
            ||||||   :::      ::::::   :::|||::::::   ::
    2149 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGT 2198

     126 aSerGluGluAspHisThrAsnAlaAlaCysPheAlaCysIleLeuLeuS 143
         :|||:::|||||||||:::::   :::   |||:::|||:::|||||||
    2199 TTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGA 2248

     143 erHisGlyGluGluAsnValIleTyrGlyLysAspGlyValThrProIle 159
         ||||||||||||||:::::::|||:::|||   :::|||   :::   :::
    2249 GCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTG 2298

     160 LysAspLeuThrAlaHisPheArgGlyAspArgCysLysThrLeuLeuGl 176
         |||   :::|||      ||||||||||||||||||:::::::|||
    2299 AAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGG 2348

     176 uLysProLysLeuPhePheIleGlnAlaCysArgGlyThrGluLeuAspA 193
          ||||||||||||||:::|||||||||||||||||||||||||||||||
    2349 AAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACT 2398

     193 spAlaIleGlnAlaAspSerGlyProIleAsnAspThrAspAlaAsnPro 209
           :::|||:::::::|||||||||      :::|||      :::
    2399 GTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGT...... 2442

     210 ArgTyrLysIleProValGluAlaAspPheLeuPheAlaTyrSerThrVa 226
           :::|||||||||||||||||||||||||||||:::|||||||||||::
    2443 ...CATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGC 2489

     226 lProGlyTyrTyrSerTrpArgSerProGlyArgGlySerTrpPheValG 243
         :|||||||||||||||||||:::          |||||||||||||:::|
    2490 ACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCC 2539

     243 lnAlaLeuCysSerIleLeuGluGluHisGlyLysGluLeuGluIleMet 259
         ||:::|||||:::::::|||:::::::::::   :::|||||:::|||
    2540 AGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATG 2589

     260 GlnIleLeuThrArgValAsnAspArgValAlaArgHisPheGluSerGl 276
         :::|||||||||||||||||   :::|||||||   :::|||||||||
    2590 CACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTT 2639

     276 nSerAspAspProHisPheHisGluLysLysGlnIleProCysValValS 293
```

```
         |||   |||      ||||||   ||||||||||||||||||:::||||
    2640 TTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTT 2689

     293 erMetLeuThrLysGluLeuTyrPhe 301
         |||||||||||||||||||||||||
    2690 CCATGCTCACAAAAGAACTCTATTTT 2715

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:AAC64043

seq_documentation_block:
ID   AAC64043 standard; DNA; 2538 BP.
XX
AC   AAC64043;
XX
DT   14-FEB-2001  (first entry)
XX
DE   pYNGCas3p10p35 insert encoding human caspase-3 and AcNPV p35 protein.
XX
KW   Human caspase-3; AcNPV p35 protein; BmNPV p10 promoter;
KW   polyhedron promoter; baculovirus vector; apoptosis inhibitor protein;
KW   recominant caspase production; insect cell; ds.
XX
OS   Chimeric - Homo sapiens.
OS   Chimeric - Autographa californica nuclear polyhedrosis virus.
OS   Chimeric - Bombyx mori nuclear polyhedrosis virus.
XX
PN   JP2000253882-A.
XX
PD   19-SEP-2000.
XX
PF   10-MAR-1999;   99JP-0062732.
XX
PR   10-MAR-1999;   99JP-0062732.
XX
PA   (KATA ) KATAKURA IND CO LTD.
XX
DR   WPI; 2000-658811/64.
XX
PT   Baculovirus vector used for the preparation of caspase protein,
PT   comprises a transfer vector plasmid containing a caspase
PT   protein-producing gene -
XX
PS   Example 1; Page 7-8; 10pp; Japanese.
XX
CC   The invention relates to a novel baculovirus vector used for the
CC   preparation of a caspase protein, and containing a recombinant transfer
CC   vector plasmid comprising a gene encoding a caspase and a gene encoding
CC   an apoptosis-inhibiting protein. The invention also encompasses the
CC   preparation of caspase protein, comprising transfecting the novel
CC   baculovirus vector into an insect or an insect cell; breeding the
CC   insect or culturing the cell; and collecting the caspase protein from
CC   the insect or insect cell. The present sequence represents the expression
CC   cassette insert from a transfer vector (pYNGCas3p10p35) produced in an
CC   exemplification of the invention. This insert comprises a mature human
CC   caspase-3 gene under the control of a polyhedron promoter, and an
CC   Autographa californica nuclear polyhedrosis virus (AcNPV) p35 gene under
CC   the control of a Bombyx mori nuclear polyhedrosis virus (BmNPV) p10
CC   promoter.
XX
SQ   Sequence 2538 BP; 865 A; 446 C; 506 G; 721 T; 0 other;


alignment_scores:
            Quality:  722.00              Length:    295
              Ratio:  3.406                Gaps:      4
 Percent Similarity:  71.864   Percent Identity: 50.169

alignment_block:
 US-09-613-508-2 x AAC64043   ..

 Align seg 1/1  to: AAC64043  from: 1  to: 2538

      28 ArgSerSerPheValProSerLeuPheSerLysLysLysLysAsnValTh 44
```

```
     |||:::|||||||||              ||||||      |||   ::
  97 CGTAACAGTTTTGTA.............ATAAAAAAACCTATAAATAGATC 134

  44 rMet.........................ArgSerIleLysThrThrA 52
     :|||                        :::||||||||||:::   :
 135 TATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGG 184

  52 rgAspArgVal..................................Pro 56
     ::   ::::::
 185 AACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTG 234

  57 ThrTyrGlnTyrAsnMetAsnPheGluLysLeuGlyLysCysIleIleIl 73
        :::|||:::|||::::::   ::::::|||   ||||||||||
 235 GACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAAT 284

  73 eAsnAsnLysAsnPheAspLysValThrGlyMetGlyValArgAsnGlyT 90
     ||||||||||||||   |||   |||||||||      |||:::||||
 285 TAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTA 334

  90 hrAspLysAspAlaGluAlaLeuPheLysCysPheArgSerLeuGlyPhe 106
     |||||   ||||||      |||   :::   ||||||:::|||   :::
 335 CAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATAT 384

 107 AspValIleValTyrAsnAspCysSerCysAlaLysMetGlnAspLeuLe 123
     :::|||      ||||||   :::      ::::::   :::|||::
 385 GAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGAT 434

 123 uLysLysAlaSerGluGluAspHisThrAsnAlaAlaCysPheAlaCysI 140
     ::::   :::|||:::||||||||:::::   :::   |||:::|||:
 435 GCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTGTTTGTG 484

 140 leLeuLeuSerHisGlyGluGluAsnValIleTyrGlyLysAspGlyVal 156
     ::|||||||||||||||||||||:::::|||:::|||   :::|||
 485 TGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCT 534

 157 ThrProIleLysAspLeuThrAlaHisPheArgGlyAspArgCysLysTh 173
     :::   :::|||   ::::|||      ||||||||||||||||||:::::
 535 GTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAG 584

 173 rLeuLeuGluLysProLysLeuPhePheIleGlnAlaCysArgGlyThrG 190
     :|||      ||||||||||||||:::||||||||||||||||||||
 585 TCTAACTGGAAAACCCAAACTTTCATTATTCAGGCCTGCCGTGGTACAG 634

 190 luLeuAspAspAlaIleGlnAlaAspSerGlyProIleAsnAspThrAsp 206
     ||||||||   :::|||:::::::||||||||      :::|||
 635 AACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCG 684

 207 AlaAsnProArgTyrLysIleProValGluAlaAspPheLeuPheAlaTy 223
     :::         :::||||||||||||||||||||||||||||||:::||||
 685 TGT.........CATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATA 725

 223 rSerThrValProGlyTyrTyrSerTrpArgSerProGlyArgGlySerT 240
     |||||||:::|||||||||||||||||||||:::         |||||||
 726 CTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCT 775

 240 rpPheValGlnAlaLeuCysSerIleLeuGluGluHisGlyLysGluLeu 256
     |||||:::|||:::|||||:::::||:::::::::::::   :::|||
 776 GGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTT 825

 257 GluIleMetGlnIleLeuThrArgValAsnAspArgValAlaArgHisPh 273
     |||:::|||:::|||||||||||||||||||   :::|||||   :::||
 826 GAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATT 875

 273 eGluSerGlnSerAspAspProHisPheHisGluLysLysGlnIleProC 290
     |||||||   |||   |||      ||||||   ||||||||||||||||
 876 TGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCAT 925

 290 ysValValSerMetLeuThrLysGluLeuTyrPhe 301
     ||:::||||||||||||||||||||||||||||||
 926 GTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 960

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAD13075

seq_documentation_block:
ID   AAD13075 standard; cDNA; 834 BP.
XX
AC   AAD13075;
XX
DT   16-OCT-2001  (first entry)
XX
DE   Mammalia caspase-3 reporter moiety cDNA.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance;
KW   caspase-3; ss.
XX
OS   Mammalia.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain -
XX
PS   Example 4; Page 131; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
CC   is also useful for developing novel assays for a wide range of
CC   post-translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is mammalia caspase-3 cDNA which is used as a reporter moiety in
CC   the exemplification of the invention.
XX
SQ   Sequence 834 BP; 274 A; 147 C; 179 G; 234 T; 0 other;


alignment_scores:
           Quality:  721.00            Length:    283
             Ratio:  3.353               Gaps:      3
 Percent Similarity:  75.972   Percent Identity: 50.883

alignment_block:
 US-09-613-508-2 x AAD13075   ..

 Align seg 1/1  to: AAD13075  from: 1  to: 834

      19 GluAspSerValAspAlaLysProAspArgSerSerPheValProSerLe 35
         |||::::|||||||:::|||      :::   ::::::   |||:::::
      13 GAAAACTCAGTGGATTCAAAATCCATTAAA...AATTTGGAACCAAAGAT 59

      35 uPheSerLysLysLysAsnValThrMetArgSerIleLysThrThrA 52
         ::::      :::::::::      :::      |||:::   ::::::
```

```
     60 CATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGT. 108

     52 rgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGly 68
                              |||:::|||::::::   ::::::|||
    109 .....................TATAAAATGGATTATCCTGAGATGGGT 135

     69 LysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGl 85
        |||||||||||||||||||||||||||   |||   |||||||||
    136 TTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGAC 185

     85 yValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheA 102
         |||::|||||||||   ||||||      |||   :::   ||||
    186 ATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCA 235

    102 rgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAlaLys 118
        ||:::|||   :::::::|||         ||||||   :::      :::
    236 GAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 285

    119 MetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAl 135
        :::   :::|||::::::   :::|||:::|||||||||::::::   ::
    286 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAG 335

    135 aCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrG 152
        :  |||:::|||:::|||||||||||||||||||||||:::::|||:::|
    336 CAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTG 385

    152 lyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheArgGly 168
        ||   :::|||   :::   :::|||   :::|||      |||||||||
    386 GAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGG 435

    169 AspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAl 185
        |||||||||:::::|||      |||||||||||||:::|||||||||
    436 GATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGC 485

    185 aCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGlyProI 202
        ||||||||||||||||||||   :::|||:::::|||||||||
    486 CTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTG 535

    202 leAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAlaAsp 218
          :::|||      :::         :::|||||||||||||||||||
    536 ATGATGACATGGCGTGT.........CATAAAATACCAGTGGAGGCCGAC 576

    219 PheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgSerPr 235
        ||||||:::|||||||||||:::|||||||||||||||||||||||:::
    577 TTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTC 626

    235 oGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGluGluH 252
              ||||||||||||:::|||:::||||||:::::::|||:::::::
    627 AAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGT 676

    252 isGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAspArg 268
        :::::   :::||||||:::|||:::||||||||||||||||||   :::
    677 ATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 726

    269 ValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisGluLy 285
        ||||||   :::||||||||||   |||   |||      ||||||   ||
    727 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAA 776

    285 sLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyrPhe 301
        |||||||||||||||:::||||||||||||||||||||||||||||||
    777 GAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 825

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT33567

seq_documentation_block:
ID   AAT33567 standard; cDNA; 1158 BP.
XX
AC   AAT33567;
XX
DT   06-DEC-1996  (first entry)
XX
DE   Pro-Yama cDNA.
XX
KW   Yama; CrmA; apoptosis; Fas receptor; T-lymphocyte; AIDS; HIV;
KW   human immunodeficiency virus; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             224..1057
FT                   /*tag= a
XX
PN   WO9625501-A2.
XX
PD   22-AUG-1996.
XX
PF   09-FEB-1996;   96WO-US01882.
XX
PR   01-JUN-1995;   95US-0457731.
PR   13-FEB-1995;   95US-0389812.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Dixit VM;
XX
DR   WPI; 1996-393402/39.
DR   P-PSDB; AAW00677.
XX
PT   New nucleic acid encoding Yama protein or CrmA mutant - useful for
PT   modulating apoptosis, maintaining T cell viability in AIDS patients
PT   and for drug screening
XX
PS   Disclosure; Fig 1; 118pp; English.
XX
CC   A cDNA clone (AAT33567) codes for pro-Yama (AAW00677), a zymogen which
CC   upon activation cleaves PARP to an 85 kDa form. Activated Yama has
CC   the ability to modulate cellular function associated with the Fas
CC   receptor pathway such as Fas-associated apoptosis. This activity is
CC   inhibited by CrmA (see also AAW00678). The cDNA was isolated from a
CC   cDNA library prepd. from human umbilical vein endothelial cells
CC   treated with tumour necrosis factor and cycloheximide. It can be
CC   used to produce Yama in a host cell for use in drug screening.
CC   Yama and anti-Yama protein nucleic acids, or antisense RNA, can be
CC   used to modulate cellular functions such as apoptosis and immune
CC   disorders mediated by the Fas pathway.
XX
SQ   Sequence 1158 BP; 328 A; 220 C; 275 G; 335 T; 0 other;

alignment_scores:
            Quality:  721.00              Length:    283
              Ratio:  3.353                 Gaps:      3
 Percent Similarity:  75.972   Percent Identity: 50.883

alignment_block:
 US-09-613-508-2 x AAT33567   ..

 Align seg 1/1  to: AAT33567  from: 1  to: 1158

     19 GluAspSerValAspAlaLysProAspArgSerSerPheValProSerLe 35
        |||:::|||||||||:::|||      :::   ::::::   |||:::::
    236 GAAAACTCAGTGGATTCAAAATCCATTAAA...AATTTGGAACCAAAGAT 282

     35 uPheSerLysLysLysLysAsnValThrMetArgSerIleLysThrThrA 52
        ::::      :::::::::      :::      |||:::   ::::::
    283 CATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGT. 331

     52 rgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGly 68
                              |||:::|||::::::   ::::::|||
    332 .....................TATAAAATGGATTATCCTGAGATGGGT 358

     69 LysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGl 85
        |||||||||||||||||||||||||||   |||   |||||||||
    359 TTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGAC 408

     85 yValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheA 102
```

```
          |||:::||||||||||   ||||||      |||   :::   ||||
    409 ATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCA 458

    102 rgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAlaLys 118
        ||:::|||   :::::|||         ||||||   :::      :::
    459 GAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 508

    119 MetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAl 135
        :::   :::|||::::::   :::|||:::||||||||||::::::   ::
    509 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAG 558

    135 aCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrG 152
        :  |||:::|||:::|||||||||||||||||||||||:::::|||:::|
    559 CAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTG 608

    152 lyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheArgGly 168
        ||   :::|||   :::   :::|||   :::|||      |||||||||
    609 GAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGG 658

    169 AspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAl 185
        |||||||||:::::|||      |||||||||||||||:::|||||||||
    659 GATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGC 708

    185 aCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGlyProI 202
        ||||||||||||||||||||   :::|||:::::::|||||||||
    709 CTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTG 758

    202 leAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAlaAsp 218
          :::|||      :::         :::||||||||||||||||||||
    759 ATGATGACATGGCGTGT.........CATAAAATACCAGTGGAGGCCGAC 799

    219 PheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgSerPr 235
        ||||||:::|||||||||||||:::|||||||||||||||||||||:::
    800 TTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTC 849

    235 oGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGluGluH 252
              ||||||||||||:::|||:::||||||:::::|||:::::::
    850 AAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGT 899

    252 isGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAspArg 268
        :::::   :::|||||:::|||:::||||||||||||||||   :::
    900 ATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 949

    269 ValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisGluLy 285
        ||||||   :::|||||||||   |||   |||      ||||||   ||
    950 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAA 999

    285 sLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyrPhe 301
        |||||||||||||||:::|||||||||||||||||||||||||||||||
   1000 GAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1048
```

seq_name: /SIDSl/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:AAV05471

```
seq_documentation_block:
ID   AAV05471 standard; DNA; 1158 BP.
XX
AC   AAV05471;
XX
DT   02-JUL-1998  (first entry)
XX
DE.  Nucleic acid encoding a protein designated YAMA.
XX
KW   CrmA; tumour necrosis factor receptor; TNF-R; U1-70; inhibition;
KW   apoptosis; activation; U1-70 path; apoptosis path; identification;
KW   modulation; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             224..1057
FT                   /*tag= a
XX
PN   JP09299077-A.
XX
PD   25-NOV-1997.
XX
PF   27-JAN-1997;   97JP-0012932.
XX
PR   09-FEB-1996;   96US-0591605.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
DR   WPI; 1998-056551/06.
DR   P-PSDB; AAW41688.
XX
PT   Controlling apoptosis by inhibiting activation of the U1-70 pathway
PT   - by introduction of nucleic acid encoding protein with CrmA
PT   biological activity into the cell
XX
PS   Disclosure; Fig 1; 48pp; Japanese.
XX
CC   The present sequence encodes a non-naturally occurring protein designated
CC   YAMA. The specification describes a novel method of controlling cell
CC   functions controlled by the tumour necrosis factor receptor (TNF-R)
CC   pathway containing U1-70. The method comprises introducing a nucleic
CC   acid molecule encoding CrmA into a cell, and culturing the cell under
CC   suitable conditions for transcription and translation of the CrmA nucleic
CC   acid. Method for preventing and inhibiting apoptosis by inhibiting
CC   activation of the U1-70 path, identifying the chemical related to the
CC   apoptosis path, and screening a candidate chemical having biological
CC   function in the apoptosis path are also described. The methods can be
CC   used to identify substances which modulate apoptosis in cells controlled
CC   by the TNF-R pathway.
XX
SQ   Sequence 1158 BP; 328 A; 220 C; 275 G; 335 T; 0 other;
```

```
alignment_scores:
            Quality:  721.00              Length:    283
              Ratio:  3.353                 Gaps:      3
 Percent Similarity:  75.972   Percent Identity: 50.883

alignment_block:
 US-09-613-508-2 x AAV05471   ..

 Align seg 1/1  to: AAV05471  from: 1  to: 1158

      19 GluAspSerValAspAlaLysProAspArgSerSerPheValProSerLe 35
         |||:::|||||||||:::|||      :::   ::::::   |||::::::
     236 GAAAACTCAGTGGATTCAAAATCCATTAAA...AATTTGGAACCAAAGAT 282

      35 uPheSerLysLysLysLysAsnValThrMetArgSerIleLysThrThrA 52
         ::::      :::::::::      :::      |||:::   ::::::
     283 CATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGT. 331

      52 rgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGly 68
                                 |||:::|||::::::   ::::::|||
     332 .....TATAAAATGGATTATCCTGAGATGGGT 358

      69 LysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGl 85
         ||||||||||||||||||||||||||||   |||   |||||||||
     359 TTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGAC 408

      85 yValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheA 102
         |||:::|||||||||   ||||||      |||   :::   ||||
     409 ATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCA 458

     102 rgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAlaLys 118
         ||:::|||   :::::|||         ||||||   :::      :::
     459 GAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 508

     119 MetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAl 135
         :::   :::|||::::::   :::|||:::||||||||||::::::   ::
     509 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAG 558

     135 aCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrG 152
```

```
     :   |||:::|||:::|||||||||||||||||||||||:::::::|||:::|
   559 CAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTG 608

   152 lyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheArgGly 168
       ||   :::|||   :::   :::|||   :::|||      |||||||||
   609 GAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGG 658

   169 AspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAl 185
       ||||||||:::::|||      |||||||||||||||:::|||||||||
   659 GATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGC 708

   185 aCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGlyProI 202
       ||||||||||||||||||||   :::|||:::::::|||||||||
   709 CTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTG 758

   202 leAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAlaAsp 218
         :::|||      :::         :::|||||||||||||||||||||
   759 ATGATGACATGGCGTGT.........CATAAAATACCAGTGGAGGCCGAC 799

   219 PheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgSerPr 235
       ||||||:::|||||||||||||:::|||||||||||||||||||||:::
   800 TTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTC 849

   235 oGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGluGluH 252
           ||||||||||||:::|||:::|||||:::::::|||:::::::
   850 AAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGT 899

   252 isGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAspArg 268
       :::::   :::|||||:::|||:::|||||||||||||||||   :::
   900 ATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 949

   269 ValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisGluLy 285
       ||||||   :::|||||||||   |||   |||      ||||||   ||
   950 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAA 999

   285 sLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyrPhe 301
       ||||||||||||||||:::|||||||||||||||||||||||||||||
  1000 GAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1048
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:AAT15277

seq_documentation_block:

```
ID   AAT15277 standard; cDNA; 1159 BP.
XX
AC   AAT15277;
XX
DT   28-OCT-1996  (first entry)
XX
DE   Human interleukin-I-converting enzyme-like apoptosis protease-4 cDNA.
XX
KW   ICE-LAP-4; interleukin-I-converting enzyme-like apoptosis protease;
KW   enzyme; Alzheimer's disease; Parkinson's disease; septic shock;
KW   head injury; rheumatoid arthritis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             115..1137
FT                   /*tag= a
XX
PN   WO9613603-A1.
XX
PD   09-MAY-1996.
XX
PF   06-JUN-1995;   95WO-US07235.
XX
PR   01-NOV-1994;   94US-0334251.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Hastings GA,  He W,  Hudson PL,  Rosen CA;
XX
DR   WPI; 1996-239509/24.
DR   P-PSDB; AAR95831.
XX
PT   Human ICE-LAP-3 and -4 DNA and protein  - useful in the diagnosis
PT   and treatment of Alzheimer's disease, Parkinson's disease,
PT   rheumatoid arthritis, septic shock and head injury
XX
PS   Claim 1; Page 44; 67pp; English.
XX
CC   This sequence was derived from a cDNA library from human endometrial
CC   tumuor and may be used for the recombinant production of the ICE-LAP-
CC   4 protein e.g. in COS cells.  The encoded protein may be used
CC   therapeutically, e.g. as an antitumor or antiviral agent and to
CC   control embryonic development and tissue homeostasis.  The protein
CC   can also be used to treat immunosuppression disorders, such as AIDS,
CC   by targeting virus infected cells for cell death.  The DNA may find
CC   use in gene therapy applications.
XX
SQ   Sequence 1159 BP; 329 A; 220 C; 276 G; 334 T; 0 other;
```

alignment_scores:

```
            Quality:  721.00            Length:    283
              Ratio:  3.353               Gaps:      3
 Percent Similarity:  75.972   Percent Identity: 50.883
```

alignment_block:

US-09-613-508-2 x AAT15277 ..

Align seg 1/1 to: AAT15277 from: 1 to: 1159

```
     19 GluAspSerValAspAlaLysProAspArgSerSerPheValProSerLe 35
        |||:::|||||||||:::|||      :::   ::::::   |||:::::
    235 GAAAACTCAGTGGATTCAAAATCCATTAAA...AATTTGGAACCAAAGAT 281

     35 uPheSerLysLysLysLysAsnValThrMetArgSerIleLysThrThrA 52
        ::::      :::::::::      :::      |||:::   ::::::
    282 CATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGT. 330

     52 rgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGly 68
                                 |||:::|||::::::   ::::::|||
    331 ......................TATAAAATGGATTATCCTGAGATGGGT 357

     69 LysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGl 85
        ||||||||||||||||||||||||||||   |||   |||||||||
    358 TTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGAC 407

     85 yValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheA 102
          |||:::|||||||||   ||||||      |||   :::   ||||
    408 ATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCA 457

    102 rgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAlaLys 118
        ||:::|||   :::::::|||         ||||||   :::      :::
    458 GAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507

    119 MetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAl 135
        :::   :::|||::::::   :::|||:::|||||||||::::::   ::
    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAG 557

    135 aCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrG 152
        :  |||:::|||:::||||||||||||||||||||||:::::|||:::|
    558 CAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTG 607

    152 lyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheArgGly 168
        ||   :::|||   :::   :::|||   :::|||      |||||||||
    608 GAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGG 657

    169 AspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAl 185
        ||||||||:::::|||      |||||||||||||||:::|||||||||
    658 GATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGC 707

    185 aCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGlyProI 202
        ||||||||||||||||||||   :::|||:::::::|||||||||
    708 CTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTG 757
```

```
202 leAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAlaAsp 218
      :::|||      :::         :::||||||||||||||||||||
758 ATGATGACATGGCGTGT.........CATAAAATACCAGTGGAGGCCGAC 798

219 PheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgSerPr 235
    ||||||:::|||||||||||||:::|||||||||||||||||||||:::
799 TTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTC 848

235 oGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGluGluH 252
          ||||||||||||:::|||:::||||||:::::|||:::::::
849 AAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGT 898

252 isGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAspArg 268
    :::::   :::||||||:::|||:::||||||||||||||||||   :::
899 ATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948

269 ValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisGluLy 285
    ||||||   :::|||||||||   |||   |||      ||||||   ||
949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAA 998

285 sLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyrPhe 301
    ||||||||||||||||:::||||||||||||||||||||||||||||||
999 GAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2002, 23:58:04 ; Search time 3311.64 Seconds
(without alignments)
7323.820 Million cell updates/sec

Title: US-09-613-508-3
Perfect score: 1159
Sequence: 1 GCACGAGCGGATGGGTGCTA.....GGGGACCTACTCTCATGCTG 1159

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1159	100.0	1159	6	AR103059	AR103059 Sequence
2	1123.8	97.0	2652	9	BC016926	BC016926 Homo sapi
3	1118	96.5	2635	6	AR172878	AR172878 Sequence
4	1118	96.5	2635	6	AX134409	AX134409 Sequence
5	1118	96.5	2635	6	AX331135	AX331135 Sequence
6	1118	96.5	2635	9	HSU13737	U13737 Human cyste
7	923.2	79.7	1001	6	AR109589	AR109589 Sequence
8	923.2	79.7	1001	9	HSU13738	U13738 Human cyste
9	921.6	79.5	1001	6	AR109591	AR109591 Sequence
10	834	72.0	834	6	AX207737	AX207737 Sequence
11	834	72.0	834	9	HSU26943	U26943 Human cyste
12	826.4	71.3	7248	6	AX207746	AX207746 Sequence
13	825	71.2	6436	6	AX207740	AX207740 Sequence
14	825	71.2	6703	6	AX207741	AX207741 Sequence
15	825	71.2	6946	6	AX207742	AX207742 Sequence
16	825	71.2	7189	6	AX207743	AX207743 Sequence
17	824.4	71.1	834	9	HSA413269	AJ413269 Homo sapi
18	690	59.5	2395	10	CGU27463	U27463 Cricetinae
19	687.4	59.3	834	4	AB029345	AB029345 Sus scrof
20	684.6	59.1	786	6	AX134427	AX134427 Sequence
21	647.2	55.8	1249	10	RNU49930	U49930 Rattus norv
22	642.4	55.4	1081	10	RNU84410	U84410 Rattus norv
23	642.4	55.4	1177	6	AR059697	AR059697 Sequence
24	642.4	55.4	1177	6	AR059698	AR059698 Sequence
25	631.8	54.5	1297	6	AR172885	AR172885 Sequence
26	631.8	54.5	1297	10	MMCASP3	Y13086 M.musculus
27	631.8	54.5	1466	10	MMU49929	U49929 Mus musculu
28	630.2	54.4	1429	10	MMU19522	U19522 Mus musculu
29	623	53.8	874	10	D86352	D86352 Mouse mRNA
30	613.6	52.9	144682	2	AC094580	AC094580 Rattus no
31	613.6	52.9	170674	2	AC096596	AC096596 Rattus no
32	578.6	49.9	884	10	RNU58656	U58656 Rattus norv
33	525.8	45.4	873	12	AF052645	AF052645 Synthetic
34	522.6	45.1	858	12	AF052647	AF052647 Synthetic
35	519.6	44.8	1673	6	AX260262	AX260262 Sequence
36	519.6	44.8	2566	6	AX260259	AX260259 Sequence
37	514.6	44.4	665	4	AF068837	AF068837 Ovis arie
c 38	500.8	43.2	131489	9	AL592104	AL592104 Human DNA
c 39	500.8	43.2	197371	2	AL355524	AL355524 Homo sapi
40	485.2	41.9	1609	10	MMU63720	U63720 Mus musculu
41	472.6	40.8	639	10	RNU34685	U34685 Rattus norv
42	438.4	37.8	3614	6	AX235203	AX235203 Sequence
43	407.8	35.2	1464	5	AF083029	AF083029 Gallus ga
c 44	306.8	26.5	169950	2	AC084049	AC084049 Homo sapi
45	306.8	26.5	207425	2	AC060790	AC060790 Homo sapi

ALIGNMENTS

RESULT 1
AR103059
LOCUS AR103059 1159 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 3 from patent US 6087150.
ACCESSION AR103059
VERSION AR103059.1 GI:12814647
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1159)
AUTHORS He,W.W., Rosen,C.A., Hudson,P.L. and Hastings,G.A.
TITLE Interleukin-1 .beta. converting enzyme like apoptosis protease-3 and 4
JOURNAL Patent: US 6087150-A 3 11-JUL-2000;
FEATURES Location/Qualifiers
source 1. .1159
/organism="unknown"
BASE COUNT 329 a 220 c 276 g 334 t
ORIGIN

Query Match 100.0%; Score 1159; DB 6; Length 1159;

```
  Best Local Similarity   100.0%;  Pred. No. 5.2e-285;
  Matches 1159;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60

Qy      61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120

Qy     121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180

Qy     181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240

Qy     241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300

Qy     301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360

Qy     361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420

Qy     421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480

Qy     481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540

Qy     541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600

Qy     601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660

Qy     661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720

Qy     721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780

Qy     781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840

Qy     841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900

Qy     901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960

Qy     961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020

Qy    1021 TCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 TCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080

Qy    1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTG 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTG 1140

Qy    1141 GGGACCTACTCTCATGCTG 1159
           ||||||||||||||||||||
Db    1141 GGGACCTACTCTCATGCTG 1159


RESULT   2
BC016926
LOCUS       BC016926               2652 bp    mRNA    linear   PRI 09-NOV-2001
DEFINITION  Homo sapiens, clone MGC:21241 IMAGE:4419175, mRNA, complete cds.
ACCESSION   BC016926
VERSION     BC016926.1  GI:16877343
KEYWORDS    MGC.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2652)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 27 Row: j Column: 23
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 14790118.
FEATURES             Location/Qualifiers
     source          1. .2652
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="MGC:21241 IMAGE:4419175"
                     /tissue_type="Lymph, lymphoma"
                     /clone_lib="NIH_MGC_85"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             228. .1061
                     /codon_start=1
                     /product="Unknown (protein for MGC:21241)"
                     /protein_id="AAH16926.1"
                     /db_xref="GI:16877344"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVEADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
BASE COUNT      824 a    466 c    576 g    786 t
ORIGIN


  Query Match           97.0%;  Score 1123.8;  DB 9;  Length 2652;
  Best Local Similarity 99.2%;  Pred. No. 5.2e-276;
```

```
  Matches 1150;  Conservative    0;  Mismatches    7;  Indels    2;  Gaps    2;

Qy       1 GCACGAGCGGATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAG 60
           ||  | ||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       6 GCGTGGGCGGATGGGTGCTATTGTGAGGCGGTTGTGGAAGAGTTTCGTGAGTGCTCGCAG 65

Qy      61 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66 CTCATACCTGTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAG 125

Qy     121 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126 GCCGTGAGGAGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCC 185

Qy     181 CTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 240
           ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     186 CTGCACCTGCCTCTTCCCCCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAAC 245

Qy     241 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     246 TCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCA 305

Qy     301 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     306 ATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTA 365

Qy     361 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     366 TGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGT 425

Qy     421 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     426 ACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGG 485

Qy     481 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     486 AATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAA 545

Qy     541 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     546 GATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATA 605

Qy     601 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     606 ATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGAT 665

Qy     661 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     666 CGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACA 725

Qy     721 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     726 GAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAA 785

Qy     781 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     786 ATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGG 845

Qy     841 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     846 CGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTAT 905

Qy     901 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     906 GCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAA 965

Qy     961 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     966 TTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTT 1025

Qy    1021 TCCATGCTCACAAAAGAACTCTATTTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1026 TCCATGCTCACAAAAGAACTCTATTTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTT 1085

Qy    1081 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTG 1140
           ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db    1086 TTTAGTTTGTATGCCAAGTGAGAAGATGGTATATTT-GGTACTGTATTTCCCTCTCATT- 1143

Qy    1141 GGGACCTACTCTCATGCTG 1159
           ||||||||||||||||||
Db    1144 TTGACCTACTCTCATGCTG 1162


RESULT   3
AR172878
LOCUS       AR172878                2635 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 3 from patent US 6303374.
ACCESSION   AR172878
VERSION     AR172878.1  GI:17912369
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2635)
  AUTHORS   Zhang,H. and Cowsert,L.M.
  TITLE     Antisense modulation of caspase 3 expression
  JOURNAL   Patent: US 6303374-A 3 16-OCT-2001;
FEATURES             Location/Qualifiers
     source          1. .2635
                     /organism="unknown"
BASE COUNT      815 a    465 c    569 g    786 t
ORIGIN


  Query Match           96.5%;  Score 1118;  DB 6;  Length 2635;
  Best Local Similarity   99.4%;  Pred. No. 1.6e-274;
  Matches 1143;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    2;

Qy      10 GATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCT 69
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12 GAGGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCT 71

Qy      70 GTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72 GTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGG 131

Qy     130 AGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     132 AGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTG 191

Qy     190 CCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGAT 249
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     192 CCTCTTCCCCCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGAT 251

Qy     250 TCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCT 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     252 TCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCT 311

Qy     310 GGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     312 GGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATA 371

Qy     370 ATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTC 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     372 ATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTC 431

Qy     430 GATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAAT 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     432 GATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAAT 491

Qy     490 GATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     492 GATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGC 551

Qy     550 AAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGA 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    552 AAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGA 611

Qy    610 ACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGA 669
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    612 ACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGA 671

Qy    670 AGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGAC 729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    672 AGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGAC 731

Qy    730 TGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTG 789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    732 TGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTG 791

Qy    790 GAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCA 849
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    792 GATGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCA 851

Qy    850 AAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAG 909
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    852 AAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAG 911

Qy    910 CTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCC 969
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    912 CTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCC 971

Qy    970 TTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTC 1029
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    972 TTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTC 1031

Qy   1030 ACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTG 1089
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1032 ACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTG 1091

Qy   1090 TATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTAC 1149
          |||||||||||||||||||||||||| |||||||||||||||||||||    ||||||
Db   1092 TATGCCAAGTGAGAAGATGGTATATTT-GGTACTGTATTTCCCTCTCATT-TTGACCTAC 1149

Qy   1150 TCTCATGCTG 1159
          ||||||||||
Db   1150 TCTCATGCTG 1159
```

```
RESULT   4
AX134409
LOCUS       AX134409                2635 bp    DNA     linear   PAT 29-MAY-2001
DEFINITION  Sequence 5 from Patent WO0129232.
ACCESSION   AX134409
VERSION     AX134409.1  GI:14271055
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2635)
  AUTHORS   Cordell,B. and Li,Y.
  TITLE     Functional cloning of genes encoding proteins/enzymes involved in
            proteolytic cleavage
  JOURNAL   Patent: WO 0129232-A 5 26-APR-2001;
            Scios Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2635
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             225. .1058
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC39521.1"
                     /db_xref="GI:14271056"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVDADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
BASE COUNT      815 a    465 c    569 g    786 t
ORIGIN
```

```
  Query Match             96.5%;  Score 1118;  DB 6;  Length 2635;
  Best Local Similarity   99.4%;  Pred. No. 1.6e-274;
  Matches 1143;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    2;

Qy     10 GATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCT 69
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     12 GAGGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCT 71

Qy     70 GTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGG 129
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     72 GTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGG 131

Qy    130 AGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTG 189
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    132 AGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTG 191

Qy    190 CCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGAT 249
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    192 CCTCTTCCCCCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGAT 251

Qy    250 TCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCT 309
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    252 TCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCT 311

Qy    310 GGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATA 369
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    312 GGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATA 371

Qy    370 ATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTC 429
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    372 ATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTC 431

Qy    430 GATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAAT 489
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    432 GATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAAT 491

Qy    490 GATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGC 549
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 GATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGC 551

Qy    550 AAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGA 609
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    552 AAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGA 611

Qy    610 ACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGA 669
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    612 ACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGA 671

Qy    670 AGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGAC 729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    672 AGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGAC 731

Qy    730 TGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTG 789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    732 TGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTG 791

Qy    790 GAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCA 849
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    792 GATGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCA 851

Qy    850 AAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAG 909
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    852 AAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAG 911

Qy    910 CTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCC 969
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    912 CTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCC 971
```

```
Qy    970 TTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTC 1029
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    972 TTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTC 1031

Qy   1030 ACAAAAGAACTCTATTTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTTAGTTTG 1089
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1032 ACAAAAGAACTCTATTTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTTAGTTTG 1091

Qy   1090 TATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTAC 1149
          |||||||||||||||||||||||||| |||||||||||||||||||||    |||||||
Db   1092 TATGCCAAGTGAGAAGATGGTATATTT-GGTACTGTATTTCCCTCTCATT-TTGACCTAC 1149

Qy   1150 TCTCATGCTG 1159
          ||||||||||
Db   1150 TCTCATGCTG 1159


RESULT   5
AX331135
LOCUS       AX331135               2635 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 1644 from Patent WO0194629.
ACCESSION   AX331135
VERSION     AX331135.1  GI:18121769
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 1644 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .2635
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      815 a    465 c    569 g    786 t
ORIGIN


  Query Match             96.5%;  Score 1118;  DB 6;  Length 2635;
  Best Local Similarity   99.4%;  Pred. No. 1.6e-274;
  Matches 1143;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    2;

Qy     10 GATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCT 69
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     12 GAGGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCT 71

Qy     70 GTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGG 129
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     72 GTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGG 131

Qy    130 AGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTG 189
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    132 AGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTG 191

Qy    190 CCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGAT 249
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    192 CCTCTTCCCCCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGAT 251

Qy    250 TCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCT 309
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    252 TCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCT 311

Qy    310 GGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATA 369
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    312 GGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATA 371

Qy    370 ATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTC 429
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    372 ATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTC 431

Qy    430 GATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAAT 489
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    432 GATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAAT 491

Qy    490 GATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGC 549
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 GATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGC 551

Qy    550 AAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGA 609
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    552 AAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGA 611

Qy    610 ACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGA 669
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    612 ACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGA 671

Qy    670 AGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGAC 729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    672 AGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGAC 731

Qy    730 TGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTG 789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    732 TGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTG 791

Qy    790 GAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCA 849
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    792 GATGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCA 851

Qy    850 AAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAG 909
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    852 AAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAG 911

Qy    910 CTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCC 969
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    912 CTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCC 971

Qy    970 TTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTC 1029
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    972 TTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTC 1031

Qy   1030 ACAAAAGAACTCTATTTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTTAGTTTG 1089
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1032 ACAAAAGAACTCTATTTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTTAGTTTG 1091

Qy   1090 TATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTAC 1149
          |||||||||||||||||||||||||| |||||||||||||||||||||    |||||||
Db   1092 TATGCCAAGTGAGAAGATGGTATATTT-GGTACTGTATTTCCCTCTCATT-TTGACCTAC 1149

Qy   1150 TCTCATGCTG 1159
          ||||||||||
Db   1150 TCTCATGCTG 1159


RESULT   6
HSU13737
LOCUS       HSU13737               2635 bp    mRNA    linear   PRI 14-APR-1995
DEFINITION  Human cysteine protease CPP32 isoform alpha mRNA, complete cds.
ACCESSION   U13737
VERSION     U13737.1  GI:561665
KEYWORDS    cysteine protease; interleukin 1-beta converting enzyme; apoptotic
            protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2635)
  AUTHORS   Fernandes-Alnemri,T., Litwack,G. and Alnemri,E.S.
  TITLE     CPP32, a novel human apoptotic protein with homology to
            Caenorhabditis elegans cell death protein Ced-3 and mammalian
            interleukin-1 beta-converting enzyme
```

```
  JOURNAL   J. Biol. Chem. 269 (49), 30761-30764 (1994)
  MEDLINE   95074098
REFERENCE   2  (bases 1 to 2635)
  AUTHORS   Fernandes-Alnemri,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-AUG-1994) Teresa Fernandes-Alnemri, Pharmacology,
            Thomas Jefferson University, Jefferson Cancer Institute, 233, S.
            Tenth Street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .2635
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="J1.1 EcoRI/XhoI fragment"
                     /cell_line="Jurkat"
                     /cell_type="T-lymphocyte"
                     /clone_lib="Human Jurkat T-lymphocyte lambda Uni-ZAP XR
                     cDNA library"
     CDS             225. .1058
                     /note="new member of the interleukin 1-beta converting
                     enzyme gene family of cysteine proteases"
                     /codon_start=1
                     /product="cysteine protease CPP32 isoform alpha"
                     /protein_id="AAA65015.1"
                     /db_xref="GI:561666"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVDADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
     polyA_site      2635
                     /note="19 A residues"
BASE COUNT      815 a    465 c    569 g    786 t
ORIGIN
```

```
  Query Match             96.5%;  Score 1118;  DB 9;  Length 2635;
  Best Local Similarity   99.4%;  Pred. No. 1.6e-274;
  Matches 1143;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    2;

Qy      10 GATGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCT 69
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12 GAGGGGTGCTATTGTGAGGCGGTTGTAGAAGAGTTTCGTGAGTGCTCGCAGCTCATACCT 71

Qy      70 GTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72 GTGGCTGTGTATCCGTGGCCACAGCTGGTTGGCGTCGCCTTGAAATCCCAGGCCGTGAGG 131

Qy     130 AGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     132 AGTTAGCGAGCCCTGCTCACACTCGGCGCTCTGGTTTTCGGTGGGTGTGCCCTGCACCTG 191

Qy     190 CCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGAT 249
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     192 CCTCTTCCCCCATTCTCATTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGAT 251

Qy     250 TCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCT 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     252 TCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCT 311

Qy     310 GGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     312 GGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATA 371

Qy     370 ATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTC 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     372 ATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTC 431

Qy     430 GATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAAT 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     432 GATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAAT 491

Qy     490 GATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     492 GATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGC 551

Qy     550 AAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGA 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     552 AAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGA 611

Qy     610 ACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGA 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     612 ACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGA 671

Qy     670 AGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGAC 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     672 AGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGAC 731

Qy     730 TGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTG 789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     732 TGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTG 791

Qy     790 GAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCA 849
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     792 GATGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCA 851

Qy     850 AAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAG 909
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     852 AAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAG 911

Qy     910 CTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCC 969
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     912 CTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCC 971

Qy     970 TTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTC 1029
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     972 TTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTC 1031

Qy    1030 ACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTG 1089
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1032 ACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTG 1091

Qy    1090 TATGCCAAGTGAGAAGATGGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTAC 1149
           |||||||||||||||||||||||||| |||||||||||||||||||   |||||||
Db    1092 TATGCCAAGTGAGAAGATGGTATATTT-GGTACTGTATTTCCCTCTCATT-TTGACCTAC 1149

Qy    1150 TCTCATGCTG 1159
           ||||||||||
Db    1150 TCTCATGCTG 1159
```

```
RESULT   7
AR109589
LOCUS       AR109589                1001 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 5 from patent US 6114132.
ACCESSION   AR109589
VERSION     AR109589.1  GI:12825865
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1001)
  AUTHORS   Desmarais,S., Gresser,M., Friesen,R., Kennedy,B., Nicholson,D.,
            Skorey,K., Ramachandran,C. and Ford-Hutchinson,A.
  TITLE     Prosphatase binding assay
  JOURNAL   Patent: US 6114132-A 5 05-SEP-2000;
FEATURES             Location/Qualifiers
     source          1. .1001
                     /organism="unknown"
BASE COUNT      310 a    171 c    227 g    293 t
ORIGIN
```

```
  Query Match             79.7%;  Score 923.2;  DB 6;  Length 1001;
  Best Local Similarity   99.5%;  Pred. No. 7.6e-225;
  Matches  947;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    2;
```

```
Qy     208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy     268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy     328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy     388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy     448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy     508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy     568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy     628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy     688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     529 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy     748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy     808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy     868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy     928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy     988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy    1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy    1108 GGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTACTCTCATGCTG 1159
           |||||||| |||||||||||||||||||||  ||||||||||||||||
Db     949 GGTATATTT-GGTACTGTATTTCCCTCTCATT-TTGACCTACTCTCATGCTG 998
```

```
RESULT   8
HSU13738
LOCUS       HSU13738                1001 bp    mRNA    linear   PRI 14-APR-1995
DEFINITION  Human cysteine protease CPP32 isoform beta mRNA, complete cds.
ACCESSION   U13738
VERSION     U13738.1  GI:561667
KEYWORDS    cysteine protease; interleukin 1-beta converting enzyme; apoptotic
            protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1001)
  AUTHORS   Fernandes-Alnemri,T., Litwack,G. and Alnemri,E.S.
  TITLE     CPP32, a novel human apoptotic protein with homology to
            Caenorhabditis elegans cell death protein Ced-3 and mammalian
            interleukin-1 beta-converting enzyme
  JOURNAL   J. Biol. Chem. 269 (49), 30761-30764 (1994)
  MEDLINE   95074098
REFERENCE   2  (bases 1 to 1001)
  AUTHORS   Fernandes-Alnemri,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-AUG-1994) Teresa Fernandes-Alnemri, Pharmacology,
            Thomas Jefferson University, Jefferson Cancer Institute, 233, S.
            Tenth Street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .1001
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="J6.1 Pst I fragment"
                     /cell_line="Jurkat"
                     /cell_type="T-lymphocyte"
                     /clone_lib="Human Jurkat T-lymphocyte lambda Uni-ZAP XR
                     cDNA library"
     5'UTR           1. .63
                     /note="An alternatively spliced isoform which lacks most
                     of the 5' nontranslated sequence"
     CDS             64. .897
                     /note="new member of the interleukin 1-beta converting
                     enzyme gene family of cysteine proteases"
                     /codon_start=1
                     /product="cysteine protease CPP32 isoform beta"
                     /protein_id="AAB60355.1"
                     /db_xref="GI:561668"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVEADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
     misc_difference 234
                     /note="differs from isoform alpha by one bp substitution
                     at position 234"
                     /replace="a"
     misc_difference 633
                     /note="differs from isoform alpha by one bp substitution
                     at position 633"
                     /replace="t"
BASE COUNT      310 a    171 c    227 g    293 t
ORIGIN
```

```
  Query Match             79.7%;  Score 923.2;  DB 9;  Length 1001;
  Best Local Similarity   99.5%;  Pred. No. 7.6e-225;
  Matches  947;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    2;

Qy     208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy     268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy     328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy     388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    529 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy    988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy   1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy   1108 GGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTACTCTCATGCTG 1159
          ||||||||| |||||||||||||||||||||||    ||||||||||||||||
Db    949 GGTATATTT-GGTACTGTATTTCCCTCTCATT-TTGACCTACTCTCATGCTG 998
```

RESULT 9
AR109591
LOCUS AR109591 1001 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 9 from patent US 6114132.
ACCESSION AR109591
VERSION AR109591.1 GI:12825867
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1001)
AUTHORS Desmarais,S., Gresser,M., Friesen,R., Kennedy,B., Nicholson,D., Skorey,K., Ramachandran,C. and Ford-Hutchinson,A.
TITLE Prosphatase binding assay
JOURNAL Patent: US 6114132-A 9 05-SEP-2000;
FEATURES Location/Qualifiers
source 1. .1001
/organism="unknown"
BASE COUNT 310 a 172 c 226 g 293 t
ORIGIN

```
Query Match           79.5%;  Score 921.6;  DB 6;  Length 1001;
Best Local Similarity 99.4%;  Pred. No. 2e-224;
Matches  946;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy    208 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 TTAATAAAGGTATCCATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAAT 108

Qy    268 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 TTGGAACCAAAGATCATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAAC 168

Qy    328 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 AGTTATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTT 228

Qy    388 CATAAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 447
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 CATAAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGG 288

Qy    448 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GAAACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

Qy    508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTT 408

Qy    568 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 TGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGAC 468

Qy    628 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 CTGAAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCC 528

Qy    688 AAACTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 747
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    529 AAACTTTTCATTATTCAGGCCTCCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGAC 588

Qy    748 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 AGTGGTGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTAT 648

Qy    808 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 GCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTC 708

Qy    868 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 ATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATT 768

Qy    928 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 CTTACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACT 828

Qy    988 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 TTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

Qy   1048 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 TATCACTAAAGAAATGGTTGGTTGGTGGTTTTTTTTAGTTTGTATGCCAAGTGAGAAGAT 948

Qy   1108 GGTATATTTGGGTACTGTATTTCCCTCTCATTGGGGACCTACTCTCATGCTG 1159
          ||||||||| |||||||||||||||||||||||    ||||||||||||||||
Db    949 GGTATATTT-GGTACTGTATTTCCCTCTCATT-TTGACCTACTCTCATGCTG 998
```

RESULT 10
AX207737
LOCUS AX207737 834 bp DNA linear PAT 31-AUG-2001
DEFINITION Sequence 35 from Patent WO0157242.

```
ACCESSION   AX207737
VERSION     AX207737.1  GI:15422421
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 834)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 35 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .834
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      274 a    147 c    179 g    234 t
ORIGIN
```

Query Match 72.0%; Score 834; DB 6; Length 834;
Best Local Similarity 100.0%; Pred. No. 4.3e-202;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     223 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATC 282
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATC 60

Qy     283 ATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGAT 342
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGAT 120

Qy     343 TATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGA 402
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGA 180

Qy     403 ATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAAC 462
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAAC 240

Qy     463 TTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 522
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 300

Qy     523 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGC 582
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGC 360

Qy     583 CATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACA 642
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACA 420

Qy     643 AACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATT 702
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 AACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATT 480

Qy     703 CAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGAT 762
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGAT 540

Qy     763 GACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 822
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 600

Qy     823 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGT 882
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGT 660

Qy     883 GCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAAC 942
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAAC 720

Qy     943 CGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAA 1002
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAA 780

Qy    1003 CAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAA 1056
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAA 834
```

```
RESULT  11
HSU26943
LOCUS       HSU26943                834 bp    mRNA    linear   PRI 29-AUG-1995
DEFINITION  Human cysteine protease Yama mRNA, complete cds.
ACCESSION   U26943
VERSION     U26943.1  GI:857568
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 834)
  AUTHORS   Tewari,M., Quan,L.T., O'Rourke,K., Desnoyers,S., Zeng,Z.,
            Beidler,D.R., Poirier,G.G., Salvesen,G.S. and Dixit,V.M.
  TITLE     Yama/CPP32 beta, a mammalian homolog of CED-3, is a
            CrmA-inhibitable protease that cleaves the death substrate
            poly(ADP-ribose) polymerase
  JOURNAL   Cell 81 (5), 801-809 (1995)
  MEDLINE   95292347
REFERENCE   2  (bases 1 to 834)
  AUTHORS   Tewari,M., O'Rourke,K., Lin,Z. and Dixit,V.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-1995) Muneesh Tewari, Pathology, University of
            Michigan, 1301 Catherine, Ann Arbor, MI 48109, USA
FEATURES             Location/Qualifiers
     source          1. .834
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="umbilical vein endothelial cells treated with
                     tumor necrosis factor and cycloheximide"
     CDS             1. .834
                     /note="cysteine protease CPP32beta; CED-3/ICE protein
                     family"
                     /codon_start=1
                     /product="Yama protein"
                     /protein_id="AAA74929.1"
                     /db_xref="GI:857569"
                     /translation="MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEM
                     GLCIIINNKNFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRD
                     VSKEDHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFII
                     QACRGTELDCGIETDSGVDDDMACHKIPVEADFLYAYSTAPGYYSWRNSKDGSWFIQS
                     LCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLTKELYFY
                     H"
BASE COUNT      274 a    147 c    179 g    234 t
ORIGIN
```

Query Match 72.0%; Score 834; DB 9; Length 834;
Best Local Similarity 100.0%; Pred. No. 4.3e-202;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     223 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATC 282
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGAGAACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATC 60

Qy     283 ATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGAT 342
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGAT 120

Qy     343 TATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGA 402
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGA 180

Qy     403 ATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAAC 462
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    181 ATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAAC 240

Qy    463 TTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATG 300

Qy    523 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGC 582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGC 360

Qy    583 CATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACA 642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACA 420

Qy    643 AACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATT 702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATT 480

Qy    703 CAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGAT 762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGAT 540

Qy    763 GACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCA 600

Qy    823 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGT 882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGT 660

Qy    883 GCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAAC 942
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAAC 720

Qy    943 CGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAA 1002
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAA 780

Qy   1003 CAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAA 1056
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAA 834
```

RESULT 12
AX207746

LOCUS	AX207746	7248 bp DNA linear PAT 31-AUG-2001
DEFINITION	Sequence 44 from Patent WO0157242.	
ACCESSION	AX207746	
VERSION	AX207746.1 GI:15422430	
KEYWORDS	.	
SOURCE	synthetic construct.	
ORGANISM	synthetic construct	
	artificial sequence.	
REFERENCE	1 (bases 1 to 7248)	
AUTHORS	Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.	
TITLE	Methods of protein destabilization and uses thereof	
JOURNAL	Patent: WO 0157242-A 44 09-AUG-2001;	
	Aurora Biosciences Corporation (US)	

FEATURES	Location/Qualifiers
source	1. .7248
	/organism="synthetic construct"
	/db_xref="taxon:32630"
	/note="pcDNA3-Ub-Met-C3 construct"

BASE COUNT 1794 a 1796 c 1826 g 1832 t
ORIGIN

Query Match 71.3%; Score 826.4; DB 6; Length 7248;
Best Local Similarity 94.9%; Pred. No. 4.2e-200;
Matches 854; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

```
Qy    172 GGGTGTGCCCTGCACCTGCCTCTTCCCGCATTCTCATTAATAAAGGTATCCATGGAGAAC 231
          | ||   ||||||||||   || |            ||  | | ||||     |||
```

```
Db   1093 GAGTCCACCCTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCACGGATCCCACATCAAC 1152

Qy    232 ACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGA 291
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   1153 ACTGAAAACTCAGTGGCCTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACATGGA 1212

Qy    292 AGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAG 351
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1213 AGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCTGAG 1272

Qy    352 ATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCT 411
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1273 ATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACATCT 1332

Qy    412 CGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATAT 471
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1333 CGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAATAT 1392

Qy    472 GAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTT 531
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1393 GAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGATGTT 1452

Qy    532 TCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAA 591
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1453 TCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAA 1512

Qy    592 GAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTC 651
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1513 GAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTC 1572

Qy    652 AGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGC 711
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1573 AGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCCTGC 1632

Qy    712 CGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCG 771
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1633 CGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATGGCG 1692

Qy    772 TGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTAT 831
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   1693 TGTCATAAAATACCAGTGGATGCCGACTTCTTGTATGCATACTCCACAGCACCTGGTTAT 1752

Qy    832 TATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTG 891
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1753 TATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTG 1812

Qy    892 AAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTG 951
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1813 AAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAGGTG 1872

Qy    952 GCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCA 1011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1873 GCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATTCCA 1932

Qy   1012 TGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACTAAAGAAATGGTTGGTTG 1071
          ||||||||||||||||||||||||||||||||||||||||  |      | ||| ||
Db   1933 TGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCACGGATCCGGGGCGTGGCTG 1992
```

RESULT 13
AX207740

LOCUS	AX207740	6436 bp DNA linear PAT 31-AUG-2001
DEFINITION	Sequence 38 from Patent WO0157242.	
ACCESSION	AX207740	
VERSION	AX207740.1 GI:15422424	
KEYWORDS	.	
SOURCE	synthetic construct.	
ORGANISM	synthetic construct	
	artificial sequence.	
REFERENCE	1 (bases 1 to 6436)	
AUTHORS	Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.	
TITLE	Methods of protein destabilization and uses thereof	
JOURNAL	Patent: WO 0157242-A 38 09-AUG-2001;	

```
          Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6436
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-1XUb-C3 construct"
BASE COUNT     1577 a   1606 c   1607 g   1646 t
ORIGIN

  Query Match             71.2%;  Score 825;  DB 6;  Length 6436;
  Best Local Similarity   100.0%;  Pred. No. 9.5e-200;
  Matches  825;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     229 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1144 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 1203

Qy     289 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 348
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1204 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 1263

Qy     349 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 408
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1264 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 1323

Qy     409 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 468
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1324 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 1383

Qy     469 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 528
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1384 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1443

Qy     529 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 588
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1444 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 1503

Qy     589 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 648
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1504 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 1563

Qy     649 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 708
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1564 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 1623

Qy     709 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 768
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1624 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 1683

Qy     769 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 828
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1684 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 1743

Qy     829 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 888
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1744 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 1803

Qy     889 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1804 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 1863

Qy     949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1008
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1864 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1923

Qy    1009 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1053
           |||||||||||||||||||||||||||||||||||||||||||||
Db    1924 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1968

RESULT  14
AX207741
```

```
LOCUS       AX207741                6703 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 39 from Patent WO0157242.
ACCESSION   AX207741
VERSION     AX207741.1  GI:15422425
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 6703)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 39 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6703
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-2XUb-C3 construct"
BASE COUNT     1653 a   1677 c   1674 g   1699 t
ORIGIN

  Query Match             71.2%;  Score 825;  DB 6;  Length 6703;
  Best Local Similarity   100.0%;  Pred. No. 9.5e-200;
  Matches  825;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     229 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1411 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 1470

Qy     289 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 348
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1471 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 1530

Qy     349 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 408
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1531 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 1590

Qy     409 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 468
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1591 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 1650

Qy     469 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 528
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1651 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1710

Qy     529 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 588
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1711 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 1770

Qy     589 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 648
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1771 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 1830

Qy     649 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 708
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1831 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 1890

Qy     709 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 768
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1891 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 1950

Qy     769 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 828
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1951 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2010

Qy     829 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 888
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2011 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 2070

Qy     889 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2071 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 2130
```

```
Qy    949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1008
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2131 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 2190

Qy   1009 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1053
          |||||||||||||||||||||||||||||||||||||||||||||
Db   2191 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 2235


RESULT 15
AX207742
LOCUS       AX207742               6946 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 40 from Patent WO0157242.
ACCESSION   AX207742
VERSION     AX207742.1  GI:15422426
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 6946)
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 40 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6946
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-3XUb-C3 construct"
BASE COUNT     1723 a   1742 c   1735 g   1746 t
ORIGIN


  Query Match             71.2%;  Score 825;  DB 6;  Length 6946;
  Best Local Similarity   100.0%;  Pred. No. 9.5e-200;
  Matches  825;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    229 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 288
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1654 AACACTGAAAACTCAGTGGATTCAAAATCCATTAAAAATTTGGAACCAAAGATCATACAT 1713

Qy    289 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 348
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1714 GGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGTTATAAAATGGATTATCCT 1773

Qy    349 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 408
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1774 GAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGACA 1833

Qy    409 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 468
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1834 TCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCAGAAACTTGAAA 1893

Qy    469 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 528
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1894 TATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATTGTGGAATTGATGCGTGAT 1953

Qy    529 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 588
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1954 GTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGTGTGCTTCTGAGCCATGGT 2013

Qy    589 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 648
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2014 GAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTT 2073

Qy    649 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 708
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2074 TTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGCC 2133

Qy    709 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 768
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2134 TGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTGATGATGACATG 2193

Qy    769 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 828
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2194 GCGTGTCATAAAATACCAGTGGAGGCCGACTTCTTGTATGCATACTCCACAGCACCTGGT 2253

Qy    829 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 888
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2254 TATTATTCTTGGCGAAATTCAAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATG 2313

Qy    889 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2314 CTGAAACAGTATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 2373

Qy    949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 1008
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2374 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAAGAAACAGATT 2433

Qy   1009 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 1053
          |||||||||||||||||||||||||||||||||||||||||||||
Db   2434 CCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTTTATCAC 2478


Search completed: August 15, 2002, 23:58:32
Job time: 6133 sec
```

```
OM of: US-09-613-508-2 to: EST:*   out_format : pfs

Date: Aug 15, 2002  8:00 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=x1h
 -Q=/cgn2_1/USPTO_spool/US09613508/runat_14082002_112647_7211/app_query.fasta_1.701
 -DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
 -MINLEN=0 -MAXLEN=2000000000 -USER=US09613508_@CGN1_1_3002
 -NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-613-508-2
Query length: 303
Database: EST:*
Database sequences: 13736207
Database length: -1841457050
Search time (sec): 2163.950000


score_list:
Sequence          Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_est1:AL542189    + 1606.00 3504.47  5.9e-186  973  ! AL542189 AL542189 LTI_FL002_PL1
gb_est2:BI821936    + 1325.00 2886.82  1.5e-151  926  ! BI821936 603039745F1 NIH_MGC_11
gb_est2:BM459840    + 1267.00 2757.64  2.4e-144 1075  ! BM459840 AGENCOURT_6422147 NIH
gb_est2:BI824214    + 1237.00 2695.55  6.8e-141  745  ! BI824214 603040584F1 NIH_MGC_11
gb_est2:BG025549    + 1156.00 2516.21  6.6e-131  830  ! BG025549 602274928F1 NIH_MGC_85
gb_est2:BF971443    + 1142.00 2487.91  2.5e-129  656  ! BF971443 602272969F1 NIH_MGC_84
gb_est1:AU126220    + 1098.00 2388.89  8.1e-124  809  ! AU126220 AU126220 NT2RP1 Homo s
gb_est1:AU130224    + 1087.00 2365.16  1.7e-122  774  ! AU130224 AU130224 NT2RP3 Homo s
gb_est1:AU134778    + 1074.50 2337.43  6.0e-121  791  ! AU134778 AU134778 PLACE1 Homo s
gb_est1:AU143139    + 1056.00 2297.75  9.7e-119  719  ! AU143139 AU143139 Y79AA1 Homo s
gb_est2:BE891855    + 1056.00 2294.66  1.4e-118  961  ! BE891855 601435409F1 NIH_MGC_72
gb_est2:BG674973    + 1010.50 2195.06  5.1e-113  917  ! BG674973 602621183F1 NCI_CGAP_S
gb_est2:BI692031    + 1006.00 2184.41  2.0e-112  984  ! BI692031 603311680F1 NCI_CGAP_M
gb_est2:BG966729    +  968.00 2103.37  6.5e-108  774  ! BG966729 602834592F1 NCI_CGAP_C
gb_est1:AI115889    +  941.00 2043.93  1.3e-104  777  ! AI115889 ue96b06.y1 Sugano mous
gb_est2:BE894817    +  919.50 1999.02  4.2e-102  621  ! BE894817 601436392F1 NIH_MGC_72
gb_est2:BE619147    +  893.50 1939.18  9.1e-99   796  ! BE619147 601472883F1 NIH_MGC_68
gb_est2:BF786558    +  891.00 1935.09  1.5e-98   697  ! BF786558 602108505F1 NCI_CGAP_K
gb_est1:AI315897    +  889.00 1929.40  3.2e-98   787  ! AI315897 uj27d07.y1 Sugano mous
gb_est2:BG339524    +  879.50 1905.89  6.5e-97  1006  ! BG339524 602437368F1 NIH_MGC_4
gb_est2:BG404254    +  824.50 1786.44  2.9e-90   870  ! BG404254 602420322F1 NIH_MGC_94
gb_est1:AU117167    +  820.00 1778.86  7.8e-90   700  ! AU117167 AU117167 HEMBA1 Homo s
gb_est2:BM150092    +  797.00 1731.18  3.5e-87   532  ! BM150092 TCAAP4Q11752 Pediatric
gb_est2:BI458770    +  796.50 1726.34  6.5e-87   756  ! BI458770 603200030F1 NIH_MGC_97
gb_est2:BM194512    +  749.00 1626.31  2.4e-81   497  ! BM194512 TCAAP2Q11752 Pediatric
gb_est2:BI401667    +  739.00 1602.52  5.2e-80   588  ! BI401667 MI-P-CP0-nvp-h-02-0-U1
gb_est2:BF540869    +  736.00 1592.54  1.9e-79   808  ! BF540869 602067587F1 NIH_MGC_58
gb_est2:BM147460    +  732.00 1589.17  2.9e-79   485  ! BM147460 TCAAP1Q11752 Pediatric
gb_est2:BF698196    +  729.50 1576.83  1.4e-78   923  ! BF698196 602125529F1 NIH_MGC_56
gb_est2:BE893968    +  723.00 1566.59  5.2e-78   630  ! BE893968 601437754F1 NIH_MGC_72
gb_est2:BI080773    +  707.00 1534.30  3.2e-76   479  ! BI080773 602878538F1 NCI_CGAP_M
gb_est2:BG264309    +  702.50 1522.69  1.4e-75   563  ! BG264309 de76e01.y1 Kirschner e
gb_est2:BM461940    +  675.00 1457.05  6.5e-72   913  ! BM461940 AGENCOURT_6418271 NIH_
gb_est2:BI652204    +  667.50 1442.03  4.5e-71   794  ! BI652204 603299029F1 NIH_CGAP_M
gb_est1:BB862907    +  664.00 1439.84  5.9e-71   473  ! BB862907 BB862907 RIKEN full-le
gb_est1:AA939788    +  652.00 1414.91  1.5e-69   412  ! AA939788 vz92g09.r1 Soares_thym
gb_est1:AI119367    +  643.00 1391.89  2.8e-68   558  ! AI119367 ue98g06.y1 Sugano mous
gb_est2:BG918158    +  636.00 1380.19  1.2e-67   394  ! BG918158 602821076F1 NCI_CGAP_M
gb_est2:BE749735    -  636.00 1379.09  1.4e-67   437  ! BE749735 200771 MARC 2PIG Sus s
gb_est1:AL534526    +  625.00 1348.11  7.6e-66   827  ! AL534526 AL534526 LTI_FL013_FBI
gb_est1:AA101428    +  617.00 1338.18  2.7e-65   402  ! AA101428 zn72e08.r1 Stratag
gb_est1:AU131461    +  605.00 1303.94  2.2e-63   840  ! AU131461 AU131461 NT2RP3 Hom
gb_est2:BE380736    +  600.00 1297.39  5.1e-63   553  ! BE380736 601270846F1 NCI_CG
gb_est2:BI657919    +  594.50 1283.52  3.0e-62   653  ! BI657919 603282147F1 NIH_CG
gb_est2:BG967610    +  588.00 1275.55  8.4e-62   360  ! BG967610 602833403F1 NCI_CG

seq_name: gb_est1:AL542189

seq_documentation_block:
LOCUS       AL542189                973 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL542189 LTI_FL002_PL1 Homo sapiens cDNA clone CS0DE008YG21 5 prime
            , mRNA sequence.
ACCESSION   AL542189
VERSION     AL542189.1  GI:12873988
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 973)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .973
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DE008YG21"
                     /clone_lib="LTI_FL002_PL1"
                     /lab_host="DH10B"
                     /note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
                     cDNA was primed with a NotI-oligo(dT) primer. Five prime
                     end enriched, double-stranded cDNA was digested with Not I
                     and cloned into the Not I and Eco RV sites of the
                     pCMVSPORT 6 vector.  Library was constructed by Life
                     Technologies. Contact : Feng Liang Life Technologies, a
                     division of Invitrogen 9800 Medical Center Drive Rockville
                     , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
                     fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      275 a    239 c    241 g    217 t      1 others
ORIGIN


alignment_scores:
           Quality: 1606.00              Length:    303
             Ratio:  5.300                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.340

alignment_block:
 US-09-613-508-2 x AL542189    ..

 Align seg 1/1  to: AL542189  from: 1  to: 973

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      35 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 84

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      85 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 134

      34 erLeuPheSerLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     135 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 184

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     185 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 234
```

```
     67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    235 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 284

     84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    285 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 334

    101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    335 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 384

    117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    385 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 434

    134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    435 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 484

    151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    485 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 534

    167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    535 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 584

    184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
        ||||||||||||||||||||||||||||||:::|||||||||||||||||
    585 AGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGG 634

    201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    635 CCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGC 684

    217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    685 TGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGA 734

    234 erProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGlu 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    735 GCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAG 784

    251 GluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAs 267
        ||||||||||||:::||||||||||||||||||||||||||||||||||
    785 GAGCACGGAAAAGACCTGGAAATCATGCAGATCCTCACCAGGGTGAATGA 834

    267 pArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisG 284
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    835 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 884

    284 luLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyr 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    885 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTAC 934

    301 PheSerGln 303
        ||||||||
    935 TTCAGTCAA 943

seq_name: gb_est2:BI821936

seq_documentation_block:
LOCUS       BI821936                926 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603039745F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5180651 5',
            mRNA sequence.
ACCESSION   BI821936
VERSION     BI821936.1  GI:15933486
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 926)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM11450  row: j  column: 12
           High quality sequence stop: 882.
FEATURES             Location/Qualifiers
     source          1. .926
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5180651"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69.  Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon  cloning).  Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was  constructed by C.
                     Gruber (Invitrogen).  Research Genetics tracking  code
                     021. Note: this is a NIH_MGC Library."
BASE COUNT      264 a    218 c    242 g    202 t
ORIGIN


alignment_scores:
            Quality: 1325.00              Length:    294
              Ratio:  4.682                Gaps:      5
 Percent Similarity:  96.259   Percent Identity: 93.878

alignment_block:
 US-09-613-508-2 x BI821936   ..

 Align seg 1/1  to: BI821936  from: 1  to: 926

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||:::|||||||||||||||||||||||||||||||||||||||
      53 ATGGCAGATGAGCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 102

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     103 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 152

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     153 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 202

      51 ThrArgAspArgValProThr.TyrGlnTyrAsnMetAsnPheGluLysL 67
         |||||||||||||||||||| ||||||||||||||||||||||||||||
     203 ACCCGGGACCGAGTGCCTACGATATCAGTACAACATGAATTTTGAAAAGC 252

      67 euGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGly 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     253 TGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGT 302

      84 MetGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCy 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     303 ATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTG 352

     100 sPheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysA 117
         |||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     353 CTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTG 402

     117 laLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsn 133
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     403 CCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAAT 452

     134 AlaAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIl 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     453 GCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAAT 502

     150 eTyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheA 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     503 TTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTA 552

     167 rgGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIle 183
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     553 GGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATT 602

     184 GlnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGl 200
         ||||||||||||||||||||||||||||||||:::|||||||||||||||
     603 CAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGG 652

     200 yProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluA 217
         ||||||||||||||||||||||||||||||||||||||      :::
     653 GCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCAGTGGAAGC 702

     217 laAspPheLeuPheAlaTyrSerThrValPro.GlyTyrTyrSerTrpAr 233
         ||||||||||||||||||||||||||||||| ||||||||||||||||||
     703 TGACTTTCCTCTTCGCCTATTCCACGGTTCCAGGGCTATTACTCGTGGAG 752

     233 gSerProGlyArg.GlySerTrpPheValGlnAlaLeu.CysSerIle.L 249
         |||||||||||||| |||||||||||||||||||||||| ||||||||| |
     753 GAGCCCAGGAAGAGGGCTCCTGGTTCGTGCAAGCCCTCTTGCTCCATTCC 802

     249 euGluGluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgVal 265
         ||||||||||||||||||   ||||||||||||||||||||||||||||||
     803 TGGAGGAGCACGGAAAAGA.CTGGAAATCATGCAGATCCTCACAAGGGTG 851

     266 AsnAspArgValAlaArgHisPheGluSerGlnSerAspAspProHisPh 282
         ||||||||||||   ||||||||||||||||||||||   |||   ||
     852 AATGACAGAGTGGC.AGGCACTTTGAGTCTCAGTCTGATGGACCCAA.TT 899

     282 eHisGluLysLysGlnIlePro 289
         |   :::||:::|||||||||
     900 CCTGAGGAAGGAACAGATCCCT 921

seq_name: gb_est2:BM459840

seq_documentation_block:
LOCUS       BM459840               1075 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6422147 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5532181
            5', mRNA sequence.
ACCESSION   BM459840
VERSION     BM459840.1  GI:18508880
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1075)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12215  row: a  column: 14
            High quality sequence stop: 545.
FEATURES             Location/Qualifiers
     source          1. .1075
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5532181"
                     /clone_lib="NIH_MGC_71"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb. "
BASE COUNT      296 a    271 c    271 g    234 t      3 others
ORIGIN


alignment_scores:
            Quality: 1267.00              Length:    301
              Ratio:  4.675                Gaps:      3
 Percent Similarity:  90.033   Percent Identity: 83.721

alignment_block:
 US-09-613-508-2 x BM459840   ..

 Align seg 1/1  to: BM459840  from: 1  to: 1075

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         |||||||||:::|||||||||||||||||||||||||||||||||||||||
      51 ATGGCAGATGAGCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 100

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 150

      34 erLeuPhe.Ser.............................LysLysLy 40
         |||||||| :::                             |||||||
     151 CCCTCTTCAGCCCCTGACTCTGGAACTTTATATTTCACCAGTAAGAAGAA 200

      40 sLysAsnValThrMetArgSerIleLysThrThrArgAspArgValProT 57
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 GAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTA 250

      57 hrTyrGlnTyrAsnMetAsnPheGluLysLeuGlyLysCysIleIleIle 73
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 CATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATA 300

      74 AsnAsnLysAsnPheAspLysValThrGlyMetGlyValArgAsnGlyTh 90
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 AACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAAC 350

      90 rAspLysAspAlaGluAlaLeuPheLysCysPheArgSerLeuGlyPheA 107
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 AGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTG 400

     107 spValIleValTyrAsnAspCysSerCysAlaLysMetGlnAspLeuLeu 123
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 ACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTT 450

     124 LysLysAlaSerGluGluAspHisThrAsnAlaAlaCysPheAlaCysIl 140
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 AAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCAT 500

     140 eLeuLeuSerHisGlyGluGluAsnValIleTyrGlyLysAspGlyValT 157
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 CCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCA 550

     157 hrProIleLysAspLeuThrAlaHisPheArgGlyAspArgCysLysThr 173
         ||||||||   ||||||||||||||||||||||||||||||||||||||||
     551 CACCAATANAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACC 600

     174 LeuLeuGluLysProLysLeuPhePheIleGlnAlaCysArgGlyThrGl 190
         ||||||||||||||||||||||:::|||||||||||||||||||:::||
     601 CTTTTAGAGAAACCCAAACTCTTTCTCATTCAGGCTTGCCGAGGNAACGA 650
```

```
     190 uLeuAspAspAlaIleGlnAlaAspSerGlyProIleAsnAspThrAspA 207
         ||||||||||:::||||||:::|||||||||   ||||||||||||||||
     651 GCTTGATGATGGCATCCAGGNCGACTCGGGGGCCATCAATGACACAGATG 700

     207 laAsnProArgTyrLysIleProValGluAlaAspPheLeuPheAlaTyr 223
         |||||||||||||   ||||||||||||:::||||||||||||   |||
     701 CTAATCCTCGATACAG.ATCCCAGTGGAAACTGACTTCCTCTTCCCCTAT 749

     224 SerThrValProGlyTyrTyrSerTrpArgSerProGlyArgGlySerTr 240
             |||   ||||||   |||||||||:::||||||:::||||||||
     750 CCACGGGTTCAAGGCTATAACTCGTGGAGGAAGCCAGGAGAAGGTTCCTG 799

     240 pPheValGlnAlaLeuCysSerIleLeuGluGluHisGlyLysGluLeuG 257
         ||||      ||||||||||||:::   ::::::||||||||:::||||
     800 GTTTGGGGCAGCCCTCTGCTCCTTCTGGAGGAAACACGGAAAAGACCTGG 849

     257 luIleMetGlnIleLeuThrArgValAsnAspArg..ValAlaArgHisP 273
         ||:::::||||||||||   ||||||||||:::|||  :::||||||   |
     850 AATTCTTGCAAATCCTCCCCAGGGTGAATGAACGAAGTTGGCCAGGACTT 899

     273 heGluSerGlnSerAspAspProHisPheHisGluLysLysGlnIlePro 289
         |||||   |||   :::||||||   |||   |||||||||      |||
     900 TTGAGGTTCAGCCGGAGGACCCCCCCTTTCCTGAAAAGAAGCCAAACCCC 949
```

seq_name: gb_est2:BI824214

seq_documentation_block:

```
LOCUS       BI824214                745 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603040584F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5181381 5',
            mRNA sequence.
ACCESSION   BI824214
VERSION     BI824214.1  GI:15935764
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 745)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11452  row: h  column: 22
            High quality sequence stop: 740.
FEATURES             Location/Qualifiers
     source          1. .745
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5181381"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69.  Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon  cloning).  Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was  constructed by C.
                     Gruber (Invitrogen).  Research Genetics tracking  code
                     021. Note: this is a NIH_MGC Library."
BASE COUNT      215 a    178 c    185 g    167 t
ORIGIN
```

alignment_scores:

```
            Quality: 1237.00              Length:    234
              Ratio:  5.286                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.145
```

alignment_block:

US-09-613-508-2 x BI824214 ..

Align seg 1/1 to: BI824214 from: 1 to: 745

```
       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      42 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 91

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      92 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 141

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     142 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 191

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     192 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 241

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     242 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 291

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     292 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 341

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     342 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 391

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     392 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 441

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         ||||||||||||||||||||||||||||||||||||||||||:::||||||
     442 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAACATGTAATT 491

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     492 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 541

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     542 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 591

     184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
         ||||||||||||||||||||||||||||||:::||||||||||||||||||
     592 AGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGG 641

     201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     642 CCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGC 691

     217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     692 TGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGA 741

     234 er 234
         ||
     742 GC 743
```

seq_name: gb_est2:BG025549

```
seq_documentation_block:
LOCUS       BG025549                 830 bp    mRNA    linear   EST 24-JAN-2001
DEFINITION  602274928F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4363038 5',
            mRNA sequence.
ACCESSION   BG025549
VERSION     BG025549.1  GI:12412258
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 830)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis Staudt, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10008  row: g  column: 07
            High quality sequence stop: 669.
FEATURES             Location/Qualifiers
     source          1. .830
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4363038"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      229 a    206 c    196 g    197 t      2 others
ORIGIN


alignment_scores:
            Quality: 1156.00            Length:    254
              Ratio:  4.797               Gaps:      7
 Percent Similarity:  94.882   Percent Identity: 90.945

alignment_block:
 US-09-613-508-2 x BG025549   ..

 Align seg 1/1  to: BG025549  from: 1  to: 830

      55 ValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGlyLysCysIl 71
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 GTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCAT 50

      71 eIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGlyValArgA 88
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 CATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAA 100

      88 snGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheArgSerLeu 104
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 ACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTG 150

     105 GlyPheAspValIleValTyrAsnAspCysSerCysAlaLysMetGlnAs 121
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 GGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGA 200

     121 pLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAlaCysPheA 138
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 TCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCG 250

     138 laCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrGlyLysAsp 154
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 CCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGAT 300

     155 GlyValThrProIleLysAspLeuThrAlaHisPheArgGlyAspArgCy 171
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 GGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATG 350

     171 sLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAlaCysArgG 188
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 CAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAG 400

     188 lyThrGluLeuAspAspAlaIleGlnAlaAspSerGlyProIleAsnAsp 204
         ||||||||||||||||:::||||||||||||||||||||||||||||||||
     401 GGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGAC 450

     205 ThrAspAlaAsnProArgTyrLysIleProValGluAlaAspPheLeuPh 221
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 ACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTT 500

     221 eAlaTyrSerThrValProGlyTyrTyrSerTrpArgSerProGlyArgG 238
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 CGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAG 550

     238 lySerTrpPheValGlnAlaLeuCysSerIleLeuGlu.GluHisGly.L 254
         ||||||||||||||||||||||||||||||||||||| |||||||| | |
     551 GCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGGAGCACGGCAA 600

     254 ysGluLeuGluIleMetGlnIleLeuThrArg.ValAsnAspArg.ValA 270
         ||:::|||||||||||||||||||||||||||    |||||||| ||||
     601 AAGACCTGGAAATCATGCAGATCCTCACCAGGGNTGAATGACAGGAGTTG 650

     270 laArgHisPheGluSerGlnSerAsp..AspProHisPheHis...GluL 285
         |||||||||||||||||||:::::  :::   ||||||      |||:
     651 CCAGGCACTTTGAGTCTCAGGCTGAATGAACNCACACTTCCCATGAGAAC 700

     285 ysLysGlnIleProCysValVal...SerMetLeuThrLysGluLeuTyr 300
         :::::|||||||||||||||||||   :::        :::||||||:::
     701 GAAGCCAGATCCCCTGTGTGGTCTTCAATGGCTCACCCAGGGAACTCTTC 750

     301 Phe 301
         |||
     751 TTT 753

seq_name: gb_est2:BF971443

seq_documentation_block:
LOCUS       BF971443                 656 bp    mRNA    linear   EST 22-JAN-2001
DEFINITION  602272969F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4361176 5',
            mRNA sequence.
ACCESSION   BF971443
VERSION     BF971443.1  GI:12338658
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 656)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10003  row: i  column: 17
            High quality sequence stop: 649.
FEATURES             Location/Qualifiers
```

```
     source          1. .656
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4361176"
                     /clone_lib="NIH_MGC_84"
                     /tissue_type="adrenal cortex carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site_1:
                     NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT
                     primed. Average insert size 1.229 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      188 a    155 c    161 g    152 t
ORIGIN


alignment_scores:
            Quality: 1142.00             Length:    219
              Ratio:  5.239               Gaps:      1
 Percent Similarity:  99.543   Percent Identity: 97.717

alignment_block:
 US-09-613-508-2 x BF971443   ..

 Align seg 1/1  to: BF971443  from: 1  to: 656

      63 AsnPheGluLysLeuGlyLysCysIleIleIleAsnAsnLysAsnPheAs 79
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
       2 AATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGA 51

      79 pLysValThrGlyMetGlyValArgAsnGlyThrAspLysAspAlaGluA 96
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
      52 TAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGG 101

      96 laLeuPheLysCysPheArgSerLeuGlyPheAspValIleValTyrAsn 112
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     102 CGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAAT 151

     113 AspCysSerCysAlaLysMetGlnAspLeuLeuLysLysAlaSerGluGl 129
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     152 GACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGA 201

     129 uAspHisThrAsnAlaAlaCysPheAlaCysIleLeuLeuSerHisGlyG 146
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     202 GGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAG 251

     146 luGluAsnValIleTyrGlyLysAspGlyValThrProIleLysAspLeu 162
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     252 AAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTG 301

     163 ThrAlaHisPheArgGlyAspArgCysLysThrLeuLeuGluLysProLy 179
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     302 ACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAA 351

     179 sLeuPhePheIleGlnAlaCysArgGlyThrGluLeuAspAspAlaIleG 196
         ||||||||||||||||||||||||||||||||||||||||||:::||||||
     352 ACTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCC 401

     196 lnAlaAspSerGlyProIleAsnAspThrAspAlaAsnProArgTyrLys 212
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     402 AGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCTCGATACAAG 451

     213 IleProValGluAlaAspPheLeuPheAlaTyrSerThrValProGlyTy 229
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     452 ATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTA 501

     229 rTyrSerTrpArgSerProGlyArgGlySer.TrpPheValGlnAlaLeu 245
         |||||||||||||||||||||||||||||| ||||||||||||||||||||
     502 TTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTTGGTTTGTGCAAGCCCTC 551

     246 CysSerIleLeuGluGluHisGlyLysGluLeuGluIleMetGlnIleLe 262
         ||||||||||||||||||||||||||:::||||||||||||||||||||||
     552 TGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT 601

     262 uThrArgValAsnAspArgValAlaArgHisPheGluSerGlnSerAspA 279
         |||||||:::||||||||||||||||||||||||||||||||||||||||:
     602 CACCAGGCTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATG 651

     279 spPro 280
         ::|||
     652 AACCA 656

seq_name: gb_est1:AU126220

seq_documentation_block:
LOCUS       AU126220                809 bp    mRNA    linear   EST 23-OCT-2000
DEFINITION  AU126220 NT2RP1 Homo sapiens cDNA clone NT2RP1000352 5', mRNA
            sequence.
ACCESSION   AU126220
VERSION     AU126220.1  GI:10950936
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 809)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Kawai,Y., Ishii,S., Saito,K.,
            Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
  TITLE     HRI human cDNA project (Ota,T., Nishikawa,T., Suzuki,Y., Kawai,Y.,
            Ishii,S., Saito,K., Nakamura,Y., Nagai,T., Sugano,S., Isogai,T.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .809
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP1000352"
                     /clone_lib="NT2RP1"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pUC19FL3; mRNA from NT2 neuronal precursor
                     cells after 48-hours retinoic acid (RA) induction"
BASE COUNT      230 a    197 c    197 g    178 t      7 others
ORIGIN


alignment_scores:
            Quality: 1098.00             Length:    231
              Ratio:  4.991               Gaps:      1
 Percent Similarity:  95.238   Percent Identity: 92.641

alignment_block:
 US-09-613-508-2 x AU126220   ..

 Align seg 1/1  to: AU126220  from: 1  to: 809

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     119 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 168

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     169 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 218

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     219 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 268

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     269 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 318

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     319 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 368

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     369 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 418

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     419 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 468

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     469 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 518

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     519 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 568

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         ||||||||||||||||||||||||||    ||||||||||||||||||||
     569 TATGGGAAAGATGGTGTCACACCAATAAANGATTTGACAGCCCACTTTAG 618

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     619 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 668

     184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
         ||||||||||||   :::   |||||||||:::||||||||||||||||||
     669 AGGCTTGCCGAAGGAACGANCTTGATGATGGCATCCAAGCCGACTCCGGG 718

     201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
         |||||||||:::   ||||||||||||||   ||||||:::|||::
     719 GCCATCAATGACCANAT.GCTAATCCTCGATACCANATCCCCTTGGAAAC 767

     217 aAspPheLeuPheAlaTyrSerThrVal.ProGlyTyrTyr 230
         :   |||   |||   ||||||   ||| ||||||:::|||
     768 TGANTTCCNCTTCCCCTATTCCCCGGTTCCCNGGTTTTTAT 808
```

seq_name: gb_est1:AU130224

seq_documentation_block:

```
LOCUS       AU130224                774 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  AU130224 NT2RP3 Homo sapiens cDNA clone NT2RP3000452 5', mRNA
            sequence.
ACCESSION   AU130224
VERSION     AU130224.1  GI:10990578
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 774)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .774
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP3000452"
                     /clone_lib="NT2RP3"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                     cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT      220 a    179 c    193 g    178 t      4 others
ORIGIN
```

alignment_scores:

```
            Quality: 1087.00             Length:    223
              Ratio:  5.032                Gaps:      2
 Percent Similarity:  96.861   Percent Identity: 96.413
```

alignment_block:

US-09-613-508-2 x AU130224 ..

Align seg 1/1 to: AU130224 from: 1 to: 774

```
       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     103 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 152

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     153 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 202

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     203 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 252

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     253 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 302

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     303 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 352

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     353 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 402

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     403 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 452

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     453 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 502

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     503 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 552

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     553 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 602

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     603 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 652

     184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAla.AspSerGl 200
         ||||||||||||||||||||||||||:::   |||||| ||||||||
```

```
     653 ANGCTTGCCGAGGGACCGAGCTTGATGATGGCATNCAGGCCCGACTCGGG 702

     200 yProIleAsnAspThrAspAlaAsnProArgTyrLysIleProVal.Glu 216
         ||||||||||||||   ||||||||||||   ||||||||||||||||
     703 GCCCATCAATGACCA.GATGCTAATCCTTGATACAAGATCCCAGTGGNAA 751

     217 AlaAspPheLeuPhe 221
         |||||||||||||||
     752 GCTGACTTTCTTTTT 766

seq_name: gb_est1:AU134778

seq_documentation_block:
LOCUS       AU134778               791 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  AU134778 PLACE1 Homo sapiens cDNA clone PLACE1000485 5', mRNA
            sequence.
ACCESSION   AU134778
VERSION     AU134778.1  GI:10995317
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 791)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .791
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="PLACE1000485"
                     /clone_lib="PLACE1"
                     /tissue_type="placenta"
                     /note="Vector: pME18SFL3"
BASE COUNT      234 a    189 c    195 g    166 t      7 others
ORIGIN


alignment_scores:
            Quality: 1074.50            Length:    222
              Ratio:  5.045              Gaps:      3
 Percent Similarity:  95.946   Percent Identity: 94.595

alignment_block:
 US-09-613-508-2 x AU134778   ..

 Align seg 1/1  to: AU134778  from: 1  to: 791

       2 AlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAlaAs 18
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      71 GCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAA 120

      18 nGluAspSerValAspAlaLysProAspArgSerSerPheValProSerL 35
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     121 TGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCC 170

      35 euPheSerLysLysLysLysAsnValThrMetArgSerIleLysThrThr 51
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     171 TCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACC 220

      52 ArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGl 68
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     221 CGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGG 270

      68 yLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyMetG 85
         |||||||||||||||||||||||||||||||||||||||||||||||:::|
     271 CAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTTTGG 320

      85 lyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysPhe 101
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     321 GCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTC 370

     102 ArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAlaLy 118
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     371 CGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAA 420

     118 sMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaA 135
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     421 GATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCG 470

     135 laCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIleTyr 151
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     471 CCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTAT 520

     152 GlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheArgGl 168
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     521 GGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGG 570

     168 yAspArgCysLysThrLeuLeuGlu.LysProLysLeuPhePheIleGln 184
         |||||||||||||||||||||||| ||||||||||||||||||||||||
     571 GGATAGATGCAAAACCCTTTTAGAAGAAACCCAAACTCTTCTTCATTCAN 620

     185 AlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGlyPr 201
         ||||||||||||||||||||||||||:::|||||||||||||||||||||
     621 GCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAAGCCGACTCCGGGGC 670

     201 oIleAsnAspThrAspAlaAsn.ProArgTyrLysIleProVal...Glu 216
         |||||||||||||||||||||| |||   ||||||   |||      :::
     671 CATCAATGACACAGATGCTAATCCCCNAATACAAAGATCCAAGTGGAAAG 720

     217 AlaAspPheLeu 220
         |||||||||
     721 CTGGACTTCCTC 732

seq_name: gb_est1:AU143139

seq_documentation_block:
LOCUS       AU143139               719 bp    mRNA    linear   EST 25-OCT-2000
DEFINITION  AU143139 Y79AA1 Homo sapiens cDNA clone Y79AA1001448 5', mRNA
            sequence.
ACCESSION   AU143139
VERSION     AU143139.1  GI:11004660
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 719)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
```

```
          Virology, Institute of Medical Science, University of Tokyo, and
          Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .719
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Y79AA1001448"
                     /clone_lib="Y79AA1"
                     /cell_type="retinoblastoma"
                     /cell_line="Y79"
                     /note="Vector: pME18SFL3"
BASE COUNT      212 a    169 c    181 g    154 t      3 others
ORIGIN


alignment_scores:
            Quality: 1056.00             Length:    219
              Ratio:  5.005               Gaps:      4
 Percent Similarity:  96.347   Percent Identity: 95.434

alignment_block:
 US-09-613-508-2 x AU143139    ..

 Align seg 1/1  to: AU143139  from: 1  to: 719

      1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
        |||||||||||||||||||||||||||||||||||||||||||||||||||
     70 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 119

     17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    120 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 169

     34 erLeuPheSerLysLysLysAsnValThrMetArgSerIleLysThr 50
        |||||||||||||||||||||||||||||||||||||||||||||||||
    170 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 219

     51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    220 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCC 269

     67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    270 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 319

     84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    320 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 369

    101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    370 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 419

    117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    420 CAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATG 469

    134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    470 CCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATT 519

    151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    520 TATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 569

    167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    570 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTC 619

    184 lnAlaCysArgGlyThr.GluLeuAspAspAlaIleGlnAla.AspSerG 200
          |||||||||||||| |||||||||||:::      ||| |||||||
    620 ANGCTTGCCGAGGGACCCGAGCTTGATGATGGCATNCANGCCCGACTCGG 669

    200 lyProIleAsn.AspThrAspAlaAsnProArgTyr.LysIleProVal 215
        |||||||||| |||||||||||||||||||||:::: ||||||||||||
    670 GGCCCATCAATGGACACAGATGCTAATCCTCGATTCCAAGATCCCAGTG 718

seq_name: gb_est2:BE891855

seq_documentation_block:
LOCUS       BE891855                961 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601435409F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920522 5',
            mRNA sequence.
ACCESSION   BE891855
VERSION     BE891855.1  GI:10351599
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 961)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9752  row: a  column: 03
            High quality sequence start: 5
            High quality sequence stop: 662.
FEATURES             Location/Qualifiers
     source          1. .961
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3920522"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      274 a    224 c    254 g    209 t
ORIGIN


alignment_scores:
            Quality: 1056.00             Length:    286
              Ratio:  4.382               Gaps:      5
 Percent Similarity:  84.266   Percent Identity: 77.972

alignment_block:
 US-09-613-508-2 x BE891855    ..

 Align seg 1/1  to: BE891855  from: 1  to: 961

      1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
        |||||||||||||||||||||||||||||||||||||||||||||||||||
     59 ATGGCAGATGATCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGC 108

     17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    109 AAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGT 158

     34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
        |||||||||||||||||||||||||||||||||||||||||||||||||
    159 CCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACC 208

     51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
        |||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
    209 ACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 258

     67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    259 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTA 308

     84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    309 TGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGC 358

    101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    359 TTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGC 408

    117 aLysMetGlnAspLeu.LeuLysLysAlaSerGluGluAspHisThrAsn 133
        |||||||||||||||| ||||||||||||||||||||||||||||||||||
    409 CAAGATGCAAGATCTGCTTAAAAAAAGCTTCTGAAGAGGACCATACAAAT 458

    134 AlaAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIl 150
        |||||||||||||||||||||||||||||||||||||||||||||||   ||
    459 GCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGT.AT 507

    150 eTyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheA 167
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    508 TTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTA 557

    167 rgGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIle 183
        ||||||||   |||||:::|||||||||||||||||||:::|||||||||
    558 GGGGGGACGA.TGCAAAAACCTTTTAGAGAAACCCAAAATCTTCTTCATT 606

    184 GlnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGl 200
        |||||||||||||||||||||||||||:::|||||||||||||||||||
    607 CAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGG 656

    200 yProIleAsnAspThrAspAla...AsnProArgTyrLysIlePro.Val 215
        ||||      :::|||:::::::   :::          |||||||| |||
    657 GCCATCAATGAGCACAGAATGCCTATCCTCCGATACAAAGATCCCAGGTG 706

    216 GluAlaAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTr 232
        |||:::   ||||||||||:::|||||||||         ||||||||||
    707 GAAGGTGGATTCCTCTTCGGCTATTCCACGGGTTCCAGGTATTACTCGTG 756

    232 pArgSerProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleL 249
        |   |||
    757 GGGGAGCCAG..................................... 766

    249 euGluGluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgVal 265
                  :::::::|||   :::   |||   |||::::::::::
    767 .............GAGAAGCTCTGGCTTGCGCAAGCCCTTGCTCATCTG 802

    266 AsnAspArgValAlaArgHisPheGluSerGlnSerAspAspProHisPh 282
        :::||||||      |||:::|||                  :::|||
    803 GAGGACAGGAAAGAGCGTAATTTC.....................AATTT 831

    282 eHis 283
        ||||
    832 CCAC 835

seq_name: gb_est2:BG674973

seq_documentation_block:
LOCUS       BG674973               917 bp    mRNA    linear   EST 01-MAY-2001
DEFINITION  602621183F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4746621 5',
            mRNA sequence.
ACCESSION   BG674973
VERSION     BG674973.1  GI:13906369
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 917)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
            Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10595  row: e  column: 22
            High quality sequence stop: 715.
FEATURES             Location/Qualifiers
     source          1. .917
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4746621"
                     /clone_lib="NCI_CGAP_Skn3"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5kb.  Library constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      263 a    251 c    202 g    201 t
ORIGIN


alignment_scores:
            Quality: 1010.50              Length:    268
              Ratio:  4.193                 Gaps:     19
 Percent Similarity:  89.925   Percent Identity: 86.940

alignment_block:
 US-09-613-508-2 x BG674973   ..

 Align seg 1/1  to: BG674973  from: 1  to: 917

     42 AsnValThrMetArgSerIleLysThrThrArgAspArgValProThrTy 58
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
      3 AATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATA 52

     58 rGlnTyrAsnMetAsnPheGluLysLeuGlyLysCysIleIleIleAsnA 75
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
     53 TCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACA 102

     75 snLysAsnPheAspLysValThrGlyMetGlyValArgAsnGlyThrAsp 91
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    103 ACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGAC 152

     92 LysAspAlaGluAlaLeuPheLysCysPheArgSerLeuGlyPheAspVa 108
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    153 AAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGT 202

    108 lIleValTyrAsnAspCysSerCysAlaLysMetGlnAspLeuLeuLysL 125
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    203 GATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAA 252

    125 ysAlaSerGluGluAspHisThrAsnAlaAlaCysPheAlaCysIleLeu 141
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    253 AAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTC 302

    142 LeuSerHisGly.GluGluAsnValIleTyrGlyLysAsp.GlyValThr 157
        |||||||||||| ||||||||||||||||||||||||||| |||||||||
    303 TTAAGCCATGGCAGAAGAAAATGTAATTTATGGGAAAGACTGGTGTCACA 352

    158 Pro.IleLysAsp.LeuThrAlaHisPhe...ArgGlyAspArgCysLys 172
        ||| :::|||||| ||||||||||||         ||||||||||||||
    353 CCAACTAAAGGACTTTGACAGCCCACTCTTAGGGCGGCGATAGATGCAAA 402

    173 ThrLeu.LeuGluLysProLysLeuPhePheIleGlnAla.CysArgGly 188
        |||||| |||||||||||||||||||||||||||||| |||||||||
    403 ACCCTTTCTAGAGAAACCCAAACTCTTCTTCATTCAGGCTATGCCGAGGG 452
```

```
    189 ThrGluLeuAsp.AspAlaIleGlnAlaAspSerGlyProIleAsnAspT 205
        ||||||||||||    :::|||||||||||||||||||||||||||||||
    453 ACCGAGCTTGATGCATGGCATCCAGGCCGACTCGGGGCCCATCAATGACA 502

    205 hrAspAlaAsnProArgTyrLysIleProValGluAlaAspPheLeuPhe 221
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    503 CAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTC 552

    222 AlaTyrSerThrVal.ProGlyTyrTyrSerTrp.ArgSerProGly.Ar 237
        |||||||||||||| ||||||||||||||    |||||||||||| ||
    553 GCCTATTCCACGGTCTCCAGGCTATTACTCGTGCGAGGAGCCCAGGAACG 602

    237 gGlySerTrp.PheValGln.AlaLeuCysSerIleLeuGluGluHisGl 253
        |||||||||| ||||||||| ||||||||||||||||||||||||||||
    603 AGGCTCCTGGTTTTGTGCAAAGCCCTCTGCTCCATCCTGGAGGAGCACGG 652

    253 yLys.GluLeuGlu.IleMetGlnIleLeuThrArgValAsn...AspAr 268
        |::: |||||||||  :::|||||||||||||||      :::      ::
    653 AAACAGAGCTGGAAACTCATGCAAATCCTCACCAGCGGTGAACTGACACA 702

    268 gValAlaArgHisPheGlu.SerGlnSerAsp.AspProHisPheHisGl 284
        :   ||||||||||||||| |||||||||:: ||||||||||||||||||
    703 GTTTGCCAGGCACTTTGAAGTCTCAGTCTGAATGACCCACACTTCCATGA 752

    284 uLysLys...GlnIleProCys 290
        |||||||   :::|||||||||
    753 GAAGAAAGCACGAATCCCCTGC 774

seq_name: gb_est2:BI692031

seq_documentation_block:
LOCUS       BI692031                 984 bp    mRNA    linear   EST 18-SEP-2001
DEFINITION  603311680F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5352042 5',
            mRNA sequence.
ACCESSION   BI692031
VERSION     BI692031.1  GI:15654660
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 984)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11894  row: o  column: 19
            High quality sequence stop: 783.
FEATURES             Location/Qualifiers
     source          1. .984
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5352042"
                     /clone_lib="NCI_CGAP_Mam6"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      249 a    271 c    288 g    176 t

ORIGIN


alignment_scores:
            Quality: 1006.00              Length:    295
              Ratio:  4.024                Gaps:      7
 Percent Similarity:  84.746   Percent Identity: 70.169

alignment_block:
 US-09-613-508-2 x BI692031   ..

 Align seg 1/1  to: BI692031  from: 1  to: 984

      1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
        |||::::|||||||||   |||      |||             ||||||::
     58 ATGACCGATGATCAGGACTGTGCTGCGGAGCTGGAAAAGGTGGATTCTTC 107

     17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheVal.Pro 33
        :::::||||||:::|||||||||||||||||||||||||||||:::::: |||
    108 CAGCGAAGACGGAGTTGACGCCAAGCCAGACCGCTCCTCTATCATCTCCT 157

     34 SerLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysTh 50
           |||:::   ||||||||||:::|||:::::          ::::::||
    158 CTATTGCTCTTGAAGAAGAAGAGAAATGCCTCTGCGGGCCCCGTCAGGAC 207

     50 rThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysL 67
        |   ||||||||||||||||||||   |||:::|||:::|||:::|||:
    208 CGGCCGGGACCGAGTGCCCACTTATCTGTACCGCATGGATTTCCAGAAGA 257

     67 euGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGly 83
        ::|||||||||||||||||||||||||||||||||||||||||:::|||||
    258 TGGGTAAATGCATCATCATAAACAACAAGAACTTCGACAAAGCGACAGGT 307

     84 MetGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCy 100
        |||   |||||||||||||||||||||||||||   ||||||||||||||
    308 ATGGACGTCCGGAATGGGACGGACAAAGATGCAGGGGCCCTCTTCAAGTG 357

    100 sPheArgSerLeuGlyPheAsp.ValIleValTyrAsnAspCysSerCys 116
        ||||:::::|||||||||::: |||   |||:::|||||||||||||||
    358 CTTCCAAAACCTGGGTTTTGAACGTAACCGTCCACAATGACTGCTCTTGT 407

    117 AlaLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAs 133
        |||||||||||||||||||||:::|||||||||||||||||||||:::||
    408 GCAAAGATGCAAGATCTGCTTAGAAAAGCCTCTGAGGAGGACCACAGCAA 457

    133 nAlaAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValI 150
        |:::|||||||||||||||:::|||||||||||||||||||||||:::::|
    458 CTCGGCCTGCTTCGCCTGCGTCCTGCTGAGCCACGGGGAAGAGGACCTGA 507

    150 leTyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPhe 166
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    508 TTTACGGGAAAGATGGCGTGACACCCATAAAGGATCTGACAGCTCATTTT 557

    167 ArgGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIl 183
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    558 AGGGGAGACCGATGCAAAACCCTGTTAGAGAAACCCAAACTCTTCTTCAT 607

    183 eGlnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerG 200
        ||||||||||||||||||||||||||||||:::|||||||||||||||||
    608 TCAGGCATGCCGAGGGACGGAGCTCGATGATGGAATCCAGGCTGACTCGG 657

    200 lyProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGlu 216
        |||||   ||||||   |||||||||||||||   |||||||||||||||
    658 GGCCATC.AACGACATTGACGCTAATCCCCGCAACAAGATCCCGGTGGAA 706

    217 AlaAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrp.A 233
        |||||||||||||||||||||||||||||||||||||||||||||||| |
    707 GCCGACTTCCTCTTTGCTTACTCCACGGTTCCAGGTTATTACTCATGGGA 756

    233 rgSerPro.GlyArgGlySerTrpPheValGlnAla......LeuCysSe 247
        ||:::||| |||:::   ||||||:::||||||::       |||
    757 GGAACCCAGGGGAAACGCTCCTGGTTGGTGCAGGGCCCTCCGGCGTGCCA 806
```

```
     247 rIleLeuGluGluHisGlyLysGluLeuGlu...IleMetGlnIleLeuT 263
                 :::   |||::::::|||:::        :::       |
     807 TCCTGAATGGAGCCTGGGCAGGGACCTCCAGAATCATGCAGAATCCGGAA 856

     263 hrArgValAsnAspArgValAla...ArgHisPheGluSerGlnSerAsp 278
         ||||||||||||||||||||:::   |||   |||:::
     857 CCAGGGTGAACGACAGGGTTGGGCCCAGGGCCTTCCGAGTCCCACGGTGA 906

     279 AspProHisPheHisGluLysLysGln 287
         :::|||:::|||   :::||||||::
     907 GAACCACGCTTCAAGAGGAAGAAGAGA 933

seq_name: gb_est2:BG966729

seq_documentation_block:
LOCUS       BG966729                774 bp    mRNA    linear   EST 12-JUN-2001
DEFINITION  602834592F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4988999 5',
            mRNA sequence.
ACCESSION   BG966729
VERSION     BG966729.1  GI:14354366
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 774)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11002  row: h  column: 24
            High quality sequence stop: 746.
FEATURES             Location/Qualifiers
     source          1. .774
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4988999"
                     /clone_lib="NCI_CGAP_Co24"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.6 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      199 a    207 c    216 g    152 t
ORIGIN


alignment_scores:
            Quality:  968.00            Length:    241
              Ratio:  4.461               Gaps:      4
 Percent Similarity:  90.041   Percent Identity: 77.178

alignment_block:
 US-09-613-508-2 x BG966729   ..

 Align seg 1/1  to: BG966729  from: 1  to: 774

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         |||:::|||||||||   |||       |||            ||||||::
      59 ATGACCGATGATCAGGACTGTGCTGCGGAGCTGGAAAAGGTGGATTCTTC 108

      17 aAsnGlu.AspSerValAspAlaLysProAspArgSerSerPheVal.Pr 33
         ::::||| |||:::|||||||||||||||||||||||||||||:::::: ||
     109 CAGCGAAGGACGGAGTTGACGCCAAGCCAGACCGCTCCTCTATCATCTCC 158

      33 oSerLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysT 50
         |   |||:::   |||||||||:::|||:::::          :::::::|
     159 TCTATTGCTCTTGAAGAAGAAGAGAAATGCCTCTGCGGGCCCCGTCAGGA 208

      50 hrThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLys 66
         ||   ||||||||||||||||||||   |||:::|||:::|||:::|||
     209 CCGGCCGGGACCGAGTGCCCACTTATCTGTACCGCATGGATTTCCAGAAG 258

      67 LeuGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGl 83
         :::|||||||||||||||||||||||||||||||||||||||||:::|||||
     259 ATGGGTAAATGCATCATCATAAACAACAAGAACTTCGACAAAGCGACAGG 308

      83 yMetGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysC 100
         ||||   |||||||||||||||||||||||||||   |||||||||||||
     309 TATGGACGTCCGGAATGGGACGGACAAAGATGCAGGGGCCCTCTTCAAGT 358

     100 ysPheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCys 116
         |||||:::::::|||||||||:::|||   |||:::|||||||||||||||
     359 GCTTCCAAAACCTGGGTTTTGAAGTAACCGTCCACAATGACTGCTCTTGT 408

     117 AlaLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAs 133
         ||||||||||||||||||||:::|||||||||||||||||||||||:::||
     409 GCAAAGATGCAAGATCTGCTTAGAAAAGCCTCTGAGGAGGACCACAGCAA 458

     133 nAlaAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValI 150
         |:::|||||||||||||||:::|||||||||||||||||||||||:::::|
     459 CTCGGCCTGCTTCGCCTGCGTCCTGCTGAGCCACGGGGAAGAGGACCTGA 508

     150 leTyr.GlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPh 166
         ||||| ||||||||||||||||||||||||||||||||||||||||||||||
     509 TTTACGGGCAAAGATGGCGTGACACCCATAAAGGATCTGACAGCTCATTT 558

     166 eArgGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheI 183
         ||||||||||||||||||||||||||||||||||||||||||||:::|||||||
     559 TAGGGGAGACCGATGCAAAACCCTGTTAGAGAAACCCAAAGTCTTCTTCA 608

     183 leGlnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSer 199
         |||||||||||||||||||||||||||||||:::|||||||||||||||
     609 TTCAGGCATGCCGAGGGACGGAGCTCGATGATGGAATCCAGGCTGACTCG 658

     200 .GlyProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValG 216
          ||||||||||||||   ||||||||||||||   ||||||||||||
     659 GGGCCCCATCAACGACATTGACGCTAATCCCCGAACAAAGATCCCGGTGG 708

     216 luAlaAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrp 232
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     709 AAGCCGACTTCCTCTTTGCTTACTCCACGGTTCCAGGTTATTACTCATGG 758

     233 ArgSerProGlyArg 237
         |||:::   |||:::
     759 AGGAACCAGGGAAAG 773

seq_name: gb_est1:AI115889

seq_documentation_block:
LOCUS       AI115889                777 bp    mRNA    linear   EST 02-SEP-1998
DEFINITION  ue96b06.y1 Sugano mouse embryo mewa Mus musculus cDNA clone
            IMAGE:1498931 5' similar to SW:ICE7_MOUSE P97864 CASPASE-7
            PRECURSOR ;, mRNA sequence.
ACCESSION   AI115889
VERSION     AI115889.1  GI:3516213
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 777)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
```

```
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:936535
            Possible reversed clone: similarity on wrong strand
            Seq primer: custom primer used
            High quality sequence stop: 507.
FEATURES             Location/Qualifiers
     source          1. .777
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1498931"
                     /clone_lib="Sugano mouse embryo mewa"
                     /dev_stage="embryo, 14 dpc"
                     /lab_host="DH10B"
                     /note="Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG);
                     Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed
                     with an oligo(dT) primer  [ATGTGGCCTTTTTTTTTTTTTTT];
                     double-stranded cDNA was ligated to a  DraIII adaptor
                     [TGTTGGCCTACTGG], digested and cloned into distinct DraIII
                     sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site
                     CACCATGTG). XhoI should be used to isolate the cDNA
                     insert.  Size selection was performed to exclude fragments
                     <1.5kb.  Library constructed by Dr. Sumio Sugano
                     (University of Tokyo Institute of Medical Science).
                     Custom primers  for sequencing: 5' end primer
                     CTTCTGCTCTAAAAGCTGCG and 3' end primer
                     CGACCTGCAGCTCGAGCACA."
BASE COUNT      203 a    210 c    207 g    154 t      3 others
ORIGIN


alignment_scores:
           Quality:  941.00              Length:    235
             Ratio:  4.546                 Gaps:      0
 Percent Similarity:  88.085   Percent Identity: 74.468

alignment_block:
 US-09-613-508-2 x AI115889   ..

 Align seg 1/1  to: AI115889  from: 1  to: 777

       1 MetAlaAspAspGlnGlyCysIleGluGluGlnGlyValGluAspSerAl 17
         |||:::|||||||||   |||     |||           ||||||::
      73 ATGACCGATGATCAGGACTGTGCTGCGGAGCTGGAAAAGGTGGATTCTTC 122

      17 aAsnGluAspSerValAspAlaLysProAspArgSerSerPheValProS 34
         ::::||||||:::||||||||||||||||||||||||||||::::::   |
     123 CAGCGAAGACGGAGTTGACGCCAAGCCAGACCGCTCCTCTATCATCTCCT 172

      34 erLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysThr 50
         ||::::::   |||||||||:::|||:::::         ::::::|||
     173 CTATTCTCTTGAAGAAGAAGAGAAATGCCTCTGCGGGCCCCGTCAGGACC 222

      51 ThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLe 67
            ||||||||||||||||||||   |||:::|||:::|||:::|||::
     223 GGCCGGGACCGAGTGCCCACTTATCTGTACCGCATGGATTTCCAGAAGAT 272

      67 uGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyM 84
         :|||||||||||||||||||||||||||||||||||||||:::|||||||
     273 GGGTAAATGCATCATCATAAACAACAAGAACTTCGACAAAGCGACAGGTA 322

      84 etGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCys 100
         ||   |||||||||||||||||||||||||||   ||||||||||||||

     323 TGGACGTCCGGAATGGGACGGACAAAGATGCAGGGGCCCTCTTCAAGTGC 372

     101 PheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAl 117
         |||:::::::|||   |||:::|||   |||:::|||||||||||||||||
     373 TTCCAAAACCTGNGTTTTGAAGTAACCGTCCACAATGACTGCTCTTGTGC 422

     117 aLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnA 134
         ||||||||||||||||||||::::|||||||||||||||||||||:::|||:
     423 AAAGATGCAAGATCTGCTTAGAAAAGCCTCTGAGGAGGACCACAGCAACT 472

     134 laAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIle 150
         ::|||||||||||||||:::|||||||||||||||||||||||:::::|||
     473 CGGCCTGCTTCGCCTGCGTCCTGCTGAGCCACGGGGAAGAGGACCTGATT 522

     151 TyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheAr 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     523 TACGGGAAAGATGGCGTGACACCCATAAAGGATCTGACAGCTCATTTTAG 572

     167 gGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleG 184
         ||||||||||||   ||||||||||||||||||||||||||||||||||
     573 GGGAGACCGATGCANAACCCTGTTAGAGAAACCCAAACTCTTCTTCATTC 622

     184 lnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGly 200
           |||||||||||||||||||||||||:::||||||||||||||||||
     623 ANGCATGCCGAGGGACGGAGCTCGATGATGGAATCCAGGCTGACTCGGGG 672

     201 ProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAl 217
         ||||||||||||   ||||||||||||||   |||   ||||||||||||
     673 CCCATCAACGACATTGACGCTAATCCCCGCAACAAGAATCCCGTGGAAGC 722

     217 aAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgS 234
         |   :::   :::||||||   ||||||   ||||||||||||   |||:
     723 CGGACTTCCTCTTGCTTACTTCACGGTTTCAGGTTATTACTCAATGAGGA 772

     234 erPro 235
         ::|||
     773 ACCCC 777
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2002, 22:16:19 ; Search time 3311.64 Seconds
(without alignments)
8663.466 Million cell updates/sec

Title: US-09-613-508-1
Perfect score: 1371
Sequence: 1 GCACGAGAAACTTTGCTGTG.....TAATTTCAGGAAAATGTATT 1371

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1371	100.0	1371	6	AR103058	AR103058 Sequence
2	1230.8	89.8	2661	9	HSU67206	U67206 Homo sapien
3	1083.8	79.1	2361	9	BC015799	BC015799 Homo sapi
4	1052.8	76.8	2335	9	HSU40281	U40281 Human cyste
5	1050.4	76.6	2309	6	AX134417	AX134417 Sequence
6	1050.4	76.6	2309	9	HSU37448	U37448 Human Mch3
7	1044.8	76.2	2314	9	HSU67320	U67320 Human Lice2
8	1007.2	73.5	2602	9	HSU67319	U67319 Human Lice2
9	886.8	64.7	912	9	HSU39613	U39613 Human cyste
10	858	62.6	2378	9	HSU37449	U37449 Human Mch3
11	679.2	49.5	2372	10	MMCASP7	Y13088 M.musculus
12	660	48.1	1063	10	AF072124	AF072124 Rattus no
13	656	47.8	2305	10	BC005428	BC005428 Mus muscu
14	655.4	47.8	2006	10	MMU67321	U67321 Mus musculu
15	632.8	46.2	1939	10	MAU47332	U47332 Mesocricetu
16	631.8	46.1	964	10	D86353	D86353 Mouse mRNA
17	364	26.5	1534	5	AB038170	AB038170 Xenopus l
18	343	25.0	155904	2	AL627395	AL627395 Homo sapi
c 19	343	25.0	174776	9	AL592546	AL592546 Human DNA
20	245.2	17.9	2395	10	CGU27463	U27463 Cricetinae
21	236.6	17.3	849	5	AB047003	AB047003 Danio rer
22	232.8	17.0	1297	6	AR172885	AR172885 Sequence
23	232.8	17.0	1297	10	MMCASP3	Y13086 M.musculus
24	232.8	17.0	1429	10	MMU19522	U19522 Mus musculu
25	232.8	17.0	1466	10	MMU49929	U49929 Mus musculu
26	232	16.9	874	10	D86352	D86352 Mouse mRNA
27	230.4	16.8	834	4	AB029345	AB029345 Sus scrof
28	229	16.7	1159	6	AR103059	AR103059 Sequence
29	229	16.7	2635	6	AR172878	AR172878 Sequence
30	229	16.7	2635	6	AX134409	AX134409 Sequence
31	229	16.7	2635	6	AX331135	AX331135 Sequence
32	229	16.7	2635	9	HSU13737	U13737 Human cyste
33	229	16.7	2652	9	BC016926	BC016926 Homo sapi
34	228.4	16.7	1249	10	RNU49930	U49930 Rattus norv
35	227.8	16.6	834	6	AX207737	AX207737 Sequence
36	227.8	16.6	834	9	HSA413269	AJ413269 Homo sapi
37	227.8	16.6	834	9	HSU26943	U26943 Human cyste
38	227.4	16.6	1001	6	AR109589	AR109589 Sequence
39	227.4	16.6	1001	9	HSU13738	U13738 Human cyste
40	226.4	16.5	6436	6	AX207740	AX207740 Sequence
41	226.4	16.5	6703	6	AX207741	AX207741 Sequence
42	226.4	16.5	6946	6	AX207742	AX207742 Sequence
43	226.4	16.5	7189	6	AX207743	AX207743 Sequence
44	226.4	16.5	7248	6	AX207746	AX207746 Sequence
45	225.8	16.5	1001	6	AR109591	AR109591 Sequence

ALIGNMENTS

RESULT 1
AR103058
LOCUS AR103058 1371 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6087150.
ACCESSION AR103058
VERSION AR103058.1 GI:12814646
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1371)
AUTHORS He,W.W., Rosen,C.A., Hudson,P.L. and Hastings,G.A.
TITLE Interleukin-1 .beta. converting enzyme like apoptosis protease-3 and 4
JOURNAL Patent: US 6087150-A 1 11-JUL-2000;
FEATURES Location/Qualifiers
source 1. .1371
/organism="unknown"
BASE COUNT . 355 a 357 c 354 g 305 t
ORIGIN

Query Match 100.0%; Score 1371; DB 6; Length 1371;

```
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1371;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCACGAGAAACTTTGCTGTGCGCGTTCTCCCGCGCGCGGGCTCAACTTTGTAGAGCGAGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GCACGAGAAACTTTGCTGTGCGCGTTCTCCCGCGCGCGGGCTCAACTTTGTAGAGCGAGG 60

Qy      61 GGCCAACTTGGCAGAGCGCGCGGCCAGCTTTGCAGAGAGCGCCCTCCAGGGACTATGCGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GGCCAACTTGGCAGAGCGCGCGGCCAGCTTTGCAGAGAGCGCCCTCCAGGGACTATGCGT 120

Qy     121 GCGGGGACACGGGTCGCTTTGGGCTCTTCCACCCCTGCGGAGCGCACTACCCCGAGCCAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCGGGGACACGGGTCGCTTTGGGCTCTTCCACCCCTGCGGAGCGCACTACCCCGAGCCAG 180

Qy     181 GGGCGGTGCAAGCCCCGCCCGGCCCTACCCAGGGCGGCTCCTCCCTCCGCAGCGCCGAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GGGCGGTGCAAGCCCCGCCCGGCCCTACCCAGGGCGGCTCCTCCCTCCGCAGCGCCGAGA 240

Qy     241 CTTTTAGTTTCGCTTTCGCTAAAGGGGCCCCAGACCCTTGCTGCGGAGCGACGGAGAGAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTTTTAGTTTCGCTTTCGCTAAAGGGGCCCCAGACCCTTGCTGCGGAGCGACGGAGAGAG 300

Qy     301 ACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGGGCTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGGGCTGT 360

Qy     361 ATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGAC 420

Qy     421 CGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCC 480

Qy     481 ATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTG 540

Qy     541 GGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGA 600

Qy     601 AACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGAC 660

Qy     661 GTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCT 720

Qy     721 GAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAA 780

Qy     781 AATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 AATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGG 840

Qy     841 GGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGA 900

Qy     901 GGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCT 960

Qy     961 AATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 AATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCA 1020

Qy    1021 GGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1021 GGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCC 1080

Qy    1081 ATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 ATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGA 1140

Qy    1141 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCA 1200

Qy    1201 GATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 GATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCA 1260

Qy    1261 GGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGC 1320

Qy    1321 TCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 TCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371


RESULT   2
HSU67206
LOCUS       HSU67206                2661 bp    mRNA    linear   PRI 03-JAN-2000
DEFINITION  Homo sapiens Lice2 alpha mRNA, complete cds.
ACCESSION   U67206
VERSION     U67206.1  GI:6649572
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2661)
  AUTHORS   Juan,T.S., McNiece,I.K., Argento,J.M., Jenkins,N.A., Gilbert,D.J.,
            Copeland,N.G. and Fletcher,F.A.
  TITLE     Identification and mapping of Casp7, a cysteine protease resembling
            CPP32 beta, interleukin-1 beta converting enzyme, and CED-3
  JOURNAL   Genomics 40 (1), 86-93 (1997)
  MEDLINE   97224489
   PUBMED   9070923
REFERENCE   2  (bases 1 to 2661)
  AUTHORS   Juan,T.S.C., McNiece,I.K., Argento,J.M., Jenkins,N.A.,
            Gilbert,D.J., Copeland,N.G. and Fletcher,F.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-AUG-1996) Developmental Hematology, Amgen, Inc., 1840
            DeHavilland Dr., Thousand Oaks, CA 91320, USA
FEATURES             Location/Qualifiers
     source          1. .2661
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             387. .1298
                     /note="like-ICE cysteine protease"
                     /codon_start=1
                     /product="Lice2 alpha"
                     /protein_id="AAF21460.1"
                     /db_xref="GI:6649573"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      740 a    569 c    642 g    709 t      1 others
ORIGIN


  Query Match           89.8%;  Score 1230.8;  DB 9;  Length 2661;
  Best Local Similarity   93.9%;  Pred. No. 0;
  Matches 1332;  Conservative    0;  Mismatches   12;  Indels   74;  Gaps    2;

Qy      27 CTCCCGCGCGCGGGCTCAACTTTGTAGAGCGAGGGGCCAACTTGGCAGAGCGCGCGGCCA 86
```

```
Db      1 CGCCCGCGCGCGGGCTCAACTTTGTAGAGCGAGGGGCCAACTTGGCAGAGCGCGCGGCCA 60

Qy     87 GCTTTGCAGAGAGCGCCCTCCAGGGACTATGCGTGCGGGGACACGGGTCGCTTTGGGCTC 146
Db     61 GCTTTGCAGAGAGCGCCCTCCAGGGACTATGCGTGCGGGGACACGGGTCGCTTTGGGCTC 120

Qy    147 TTCCACCCCTGCGGAGCGCACTACCCCGAGCCAGGGGCGGTGCAAGCCCCGCCCGGCCCT 206
Db    121 TTCCACCCCTGCGGAGCGCACTATCCCGAGCCAGGGGCGGTGCAAGCCCCGCCCGGCCCT 180

Qy    207 ACCCAGGGCGGCTCCTCCCTCCGCAGCGCCGAGACTTTTAGTTTCGCTTTCGCTAAAGGG 266
Db    181 ACCCAGGGCGGCTCCTCCCTCCGCAGCGCCGAGACTTTTAGTTTCGCTTTCGCTAAAGGG 240

Qy    267 GCCCCAGACCCTTGCTGCGGAGCGACGGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACC 326
Db    241 GCCCCAGACCCTTGCTGCGGAGCGACGGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACC 300

Qy    327 GCCGTGGGAACGATGGCAGATG---------------------------------- 348
Db    301 GCCGTGGGAACGGCAGGAAGTGGCACTTGGAAAAGAACACCAGCTGCGGTGGTAGCAGTG 360

Qy    349 ----------------------------------ATTCAGGGCTGTATTGAAGAGCAGG 373
Db    361 GGATTTGTGCTTCTTATGTTACCCAGATGGCAGATGATCAGGGCTGTATTGAAGAGCAGG 420

Qy    374 GGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTG 433
Db    421 GGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTG 480

Qy    434 TACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCC 493
Db    481 TACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCC 540

Qy    494 GGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCA 553
Db    541 GGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCA 600

Qy    554 TAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACA 613
Db    601 TAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACA 660

Qy    614 AAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATA 673
Db    661 AAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATA 720

Qy    674 ATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATA 733
Db    721 ATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATA 780

Qy    734 CAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATG 793
Db    781 CAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATG 840

Qy    794 GGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCA 853
Db    841 GGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCA 900

Qy    854 AAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTG 913
Db    901 AAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTG 960

Qy    914 ATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCTCGATACA 973
Db    961 ATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCTCGATACA 1020

Qy    974 AGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGT 1033
Db   1021 AGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCAT 1080

Qy   1034 GGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGC 1093
```

```
Db   1081 GGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGC 1140

Qy   1094 ACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGG 1153
Db   1141 ACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCAGGGTGAATGACAGAGTTGCCAGG 1199

Qy   1154 CACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTG 1213
Db   1200 CACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTG 1259

Qy   1214 GTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCAGGGGTACATTCTA 1273
Db   1260 GTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCAGGGGTACATTCTA 1319

Qy   1274 GCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGCTCTTGAAATATTC 1333
Db   1320 GCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGCTCTTGAAATATTC 1379

Qy   1334 AGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
Db   1380 AGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1417
```

RESULT 3
BC015799

LOCUS BC015799 2361 bp mRNA linear PRI 29-OCT-2001

DEFINITION Homo sapiens, caspase 7, apoptosis-related cysteine protease, clone MGC:9176 IMAGE:3921455, mRNA, complete cds.

ACCESSION BC015799

VERSION BC015799.1 GI:16041818

KEYWORDS MGC.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2361)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (09-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov

COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 15 Row: n Column: 12
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 15718696.

FEATURES Location/Qualifiers

source 1. .2361
/organism="Homo sapiens"
/db_xref="LocusID:840"
/db_xref="taxon:9606"
/clone="MGC:9176 IMAGE:3921455"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NIH_MGC_72"

```
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             79. .990
                     /codon_start=1
                     /product="caspase 7, apoptosis-related cysteine protease"
                     /protein_id="AAH15799.1"
                     /db_xref="GI:16041819"
                     /translation="MADEQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      704 a    464 c    545 g    648 t
ORIGIN


  Query Match             79.1%;  Score 1083.8;  DB 9;  Length 2361;
  Best Local Similarity   99.6%;  Pred. No. 1.9e-287;
  Matches 1107;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    2;

Qy     261 AAAGGGGCCCCAGACCCTTGCTGCGGAGCGACGGAGAGAGACTGTGCCAGTCCCAGCCGC 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 AAAGGGGCCCCAGACCCTTGCTGCGGAGCGACGGAGAGAGACTGTGCCAGTCCCAGCCGC 60

Qy     321 CCTACCGCCGTGGGAACGATGGCAGATGATTCAGGGCTGTATTGAAGAGCAGGGGGTTGA 380
           |||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db      61 CCTACCGCCGTGGGAACGATGGCAGATGA-GCAGGGCTGTATTGAAGAGCAGGGGGTTGA 119

Qy     381 GGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTC 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     120 GGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTC 179

Qy     441 CCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCG 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180 CCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCG 239

Qy     501 AGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAA 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     240 AGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAA 299

Qy     561 CAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGC 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 CAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGC 359

Qy     621 CGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTG 680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 CGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTG 419

Qy     681 CTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGC 740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     420 CTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGC 479

Qy     741 CGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAAATGTAATTTATGGGAAAGA 800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     480 CGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGA 539

Qy     801 TGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCT 860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     540 TGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCT 599

Qy     861 TTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGG 920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 TTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGG 659

Qy     921 CATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCC 980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     660 CATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCC 719

Qy     981 AGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAG 1040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     720 AGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAG 779

Qy    1041 CCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAA 1100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     780 CCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAA 839

Qy    1101 AGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTG 1160
           |||||||||||||||||  ||||| ||||||||||||||||||||||||||||||||||||
Db     840 AGACCTGGAAATCATGCAGATCCT-CACCAGGGTGAATGACAGAGTTGCCAGGCACTTTG 898

Qy    1161 AGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCA 1220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     899 AGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCA 958

Qy    1221 TGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCAGGGGTACATTCTAGCTGAGA 1280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     959 TGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCAGGGGTACATTCTAGCTGAGA 1018

Qy    1281 AGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGCTCTTGAAATATTCAGAAATT 1340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1019 AGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGCTCTTGAAATATTCAGAAATT 1078

Qy    1341 CTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
           |||||||||||||||||||||||||||||||
Db    1079 CTCCAGGATTTTAATTTCAGGAAAATGTATT 1109


RESULT   4
HSU40281
LOCUS       HSU40281                2335 bp    mRNA    linear   PRI 27-JAN-1996
DEFINITION  Human cysteine protease CMH-1 mRNA, complete cds.
ACCESSION   U40281
VERSION     U40281.1  GI:1167554
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2335)
  AUTHORS   Lippke,J.A., Gu,Y., Sarnecki,C., Caron,P.R. and Su,M.S.
  TITLE     Identification and characterization of CPP32/Mch2 homolog 1, a
            novel cysteine protease similar to CPP32
  JOURNAL   J. Biol. Chem. 271 (4), 1825-1828 (1996)
  MEDLINE   96147144
REFERENCE   2  (bases 1 to 2335)
  AUTHORS   Caron,P.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-NOV-1995) Paul R. Caron, Vertex Pharmaceuticals, 40
            Allston St., Cambridge, MA 02139, USA
FEATURES             Location/Qualifiers
     source          1. .2335
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="IMAGE consortium clone 72778"
     CDS             48. .959
                     /function="cysteine protease"
                     /note="similar to CPP32 and to MCH2; related to Ced3 and
                     interleukin-1 beta converting enzyme"
                     /codon_start=1
                     /product="CMH-1"
                     /protein_id="AAC50352.1"
                     /db_xref="GI:1167555"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      700 a    454 c    535 g    646 t
ORIGIN


  Query Match             76.8%;  Score 1052.8;  DB 9;  Length 2335;
  Best Local Similarity   99.6%;  Pred. No. 6.5e-279;
  Matches 1076;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    2;
```

```
Qy    292 CGGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATT 351
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      1 CGGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGA-T 59

Qy    352 CAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCT 411
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     60 CAGGGCTGTATTGAGGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCT 119

Qy    412 AAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACC 471
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 AAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACC 179

Qy    472 ATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTT 531
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 ATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTT 239

Qy    532 GAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATG 591
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 GAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATG 299

Qy    592 GGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTG 651
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 GGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTG 359

Qy    652 GGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAA 711
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 GGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAA 419

Qy    712 AAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCAT 771
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 AAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCAT 479

Qy    772 GGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCC 831
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 GGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCC 539

Qy    832 CACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAG 891
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 CACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAG 599

Qy    892 GCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGAC 951
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 GCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGAC 659

Qy    952 ACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCC 1011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 ACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCC 719

Qy   1012 ACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCC 1071
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 ACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCC 779

Qy   1072 CTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAG 1131
          |||||||||||||||||||||||||||||||||||||||||||||| ||||| ||||||
Db    780 CTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCAG 838

Qy   1132 GGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGA 1191
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    839 GGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGA 898

Qy   1192 GAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATA 1251
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    899 GAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATA 958

Qy   1252 GCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATT 1311
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    959 GCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATT 1018

Qy   1312 TTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1019 TTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1078
```

```
RESULT   5
AX134417
LOCUS       AX134417               2309 bp    DNA     linear   PAT 29-MAY-2001
DEFINITION  Sequence 13 from Patent WO0129232.
ACCESSION   AX134417
VERSION     AX134417.1  GI:14271063
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2309)
  AUTHORS   Cordell,B. and Li,Y.
  TITLE     Functional cloning of genes encoding proteins/enzymes involved in
            proteolytic cleavage
  JOURNAL   Patent: WO 0129232-A 13 26-APR-2001;
            Scios Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2309
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             44. .955
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC39525.1"
                     /db_xref="GI:14271064"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      680 a    453 c    531 g    645 t
ORIGIN
```

Query Match 76.6%; Score 1050.4; DB 6; Length 2309;
Best Local Similarity 99.7%; Pred. No. 3e-278;
Matches 1073; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```
Qy    296 GAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGG 355
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db      1 GAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGA-TCAGG 59

Qy    356 GCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGC 415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 GCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGC 119

Qy    416 CAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGC 475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 CAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGC 179

Qy    476 GATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAA 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 GATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAA 239

Qy    536 AGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCG 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 AGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCG 299

Qy    596 TTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTT 655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 TTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTT 359

Qy    656 TTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAG 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 TTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAG 419

Qy    716 CTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAG 775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 CTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAG 479
```

```
Qy    776 AAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACT 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 AAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACT 539

Qy    836 TTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTT 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 TTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTT 599

Qy    896 GCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAG 955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 GCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAG 659

Qy    956 ATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGG 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 ATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGG 719

Qy   1016 TTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCT 1075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 TTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCT 779

Qy   1076 GCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTG 1135
          |||||||||||||||||||||||||||||||||||| ||||| ||||||||||
Db    780 GCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCAGGGTG 838

Qy   1136 AATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAG 1195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    839 AATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAG 898

Qy   1196 AAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCA 1255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    899 AAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCA 958

Qy   1256 TATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTT 1315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    959 TATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTT 1018

Qy   1316 TATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1019 TATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1074
```

```
RESULT   6
HSU37448
LOCUS       HSU37448                 2309 bp    mRNA    linear   PRI 13-DEC-1995
DEFINITION  Human Mch3 isoform alpha (Mch3) mRNA, complete cds.
ACCESSION   U37448
VERSION     U37448.1  GI:1117846
KEYWORDS    Mch3 isoform alpha.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2309)
  AUTHORS   Fernandes-Alnemri,T., Litwack,G. and Alnemri,E.S.
  TITLE     CPP32, a novel human apoptotic protein with homology to
            Caenorhabditis elegans cell death protein Ced-3 and mammalian
            interleukin-1 beta-converting enzyme
  JOURNAL   J. Biol. Chem. 269 (49), 30761-30764 (1994)
  MEDLINE   95074098
REFERENCE   2  (bases 1 to 2309)
  AUTHORS   Fernandes-Alnemri,T., Takahashi,A., Armstrong,R., Krebs,J.,
            Fritz,L., Tomaselli,K.J., Wang,L., Yu,Z., Croce,C.M., Salveson,G.,
            Earnshaw,W.C., Litwack,G. and Alnemri,E.S.
  TITLE     Mch3, a novel human apoptotic cysteine protease highly related to
            CPP32
  JOURNAL   Cancer Res. 55 (24), 6045-6052 (1995)
  MEDLINE   96105019
REFERENCE   3  (bases 1 to 2309)
  AUTHORS   Alnemri,E.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-1995) Emad S. Alnemri, Pharmacology, Thomas
            Jefferson University, Jefferson Cancer Institute, 233, S. Tenth
            Street, Philadelphia, PA 19107, USA
```

```
FEATURES             Location/Qualifiers
     source          1. .2309
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="T50-8A.1"
                     /cell_line="Jurkat"
                     /cell_type="T-lymphocyte"
     gene            1. .2309
                     /gene="Mch3"
     CDS             44. .955
                     /gene="Mch3"
                     /function="cysteine protease"
                     /note="new member of ICE-like cysteine protease family"
                     /codon_start=1
                     /evidence=experimental
                     /product="Mch3 isoform alpha"
                     /protein_id="AAC50303.1"
                     /db_xref="GI:1117847"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
     polyA_site      2309
                     /gene="Mch3"
BASE COUNT      680 a    453 c    531 g    645 t
ORIGIN
```

```
  Query Match           76.6%;  Score 1050.4;  DB 9;  Length 2309;
  Best Local Similarity  99.7%;  Pred. No. 3e-278;
  Matches 1073;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy    296 GAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGG 355
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db      1 GAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGA-TCAGG 59

Qy    356 GCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGC 415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 GCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGC 119

Qy    416 CAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGC 475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 CAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGC 179

Qy    476 GATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAA 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 GATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAA 239

Qy    536 AGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCG 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 AGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCG 299

Qy    596 TTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTT 655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 TTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTT 359

Qy    656 TTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAG 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 TTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAG 419

Qy    716 CTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAG 775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 CTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAG 479

Qy    776 AAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACT 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 AAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACT 539

Qy    836 TTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTT 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 TTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTT 599
```

```
Qy    896 GCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAG 955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 GCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAG 659

Qy    956 ATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGG 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 ATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGG 719

Qy   1016 TTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCT 1075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 TTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCT 779

Qy   1076 GCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTG 1135
          |||||||||||||||||||||||||||||||||||||| ||||| ||||||||||
Db    780 GCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCAGGGTG 838

Qy   1136 AATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAG 1195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    839 AATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAG 898

Qy   1196 AAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCA 1255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    899 AAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCA 958

Qy   1256 TATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTT 1315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    959 TATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTT 1018

Qy   1316 TATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1019 TATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1074
```

```
RESULT   7
HSU67320
LOCUS       HSU67320                2314 bp    mRNA    linear   PRI 19-MAR-1997
DEFINITION  Human Lice2 gamma cysteine protease mRNA, complete cds.
ACCESSION   U67320
VERSION     U67320.1  GI:1894914
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2314)
  AUTHORS   Juan,T.S., McNiece,I.K., Argento,J.M., Jenkins,N.A., Gilbert,D.J.,
            Copeland,N.G. and Fletcher,F.A.
  TITLE     Identification and mapping of Casp7, a cysteine protease resembling
            CPP32 beta, interleukin-1 beta converting enzyme, and CED-3
  JOURNAL   Genomics 40 (1), 86-93 (1997)
  MEDLINE   97224489
REFERENCE   2  (bases 1 to 2314)
  AUTHORS   Juan,T.S.C., McNiece,I.K., Argento,J.M., Jenkins,N.A.,
            Gilbert,D.J., Copeland,N.G. and Fletcher,F.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-1996) Developmental Hematology, Amgen, Inc., 1840
            DeHavilland Dr., Thousand Oaks, CA 91320, USA
FEATURES             Location/Qualifiers
     source          1. .2314
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             40. .951
                     /note="similar to ICE and CED-3 cysteine protease"
                     /codon_start=1
                     /product="Lice2 gamma cysteine protease"
                     /protein_id="AAC51153.1"
                     /db_xref="GI:1894915"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      680 a    459 c    526 g    648 t      1 others
ORIGIN
```

```
  Query Match           76.2%;  Score 1044.8;  DB 9;  Length 2314;
  Best Local Similarity  99.6%;  Pred. No. 1e-276;
  Matches 1068;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    2;

Qy    300 GACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGGGCTG 359
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db      1 GACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGA-TCAGGGCTG 59

Qy    360 TATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGA 419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 TATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGA 119

Qy    420 CCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATC 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 CCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATC 179

Qy    480 CATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 CATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCT 239

Qy    540 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCG 599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 GGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCG 299

Qy    600 AAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGA 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 AAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGA 359

Qy    660 CGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTC 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 CGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTC 419

Qy    720 TGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGA 779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 TGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGA 479

Qy    780 AAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 AAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAG 539

Qy    840 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCG 899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 GGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCG 599

Qy    900 AGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGC 959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 AGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGC 659

Qy    960 TAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCC 1019
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 TAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCC 719

Qy   1020 AGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTC 1079
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    720 AGGCTATTACTCATGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTC 779

Qy   1080 CATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATG 1139
          ||||||||||||||||||||||||||||||||||||| ||||| ||||||||||||||||
Db    780 CATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCAGGGTGAATG 838

Qy   1140 ACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGC 1199
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    839 ACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGC 898

Qy   1200 AGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATC 1259
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    899 AGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATC 958
```

```
Qy   1260 AGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATG 1319
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    959 AGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATG 1018

Qy   1320 CTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1019 CTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1070


RESULT   8
HSU67319
LOCUS       HSU67319                2602 bp    mRNA    linear   PRI 19-MAR-1997
DEFINITION  Human Lice2 beta cysteine protease mRNA, complete cds.
ACCESSION   U67319
VERSION     U67319.1  GI:1894912
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2602)
  AUTHORS   Juan,T.S., McNiece,I.K., Argento,J.M., Jenkins,N.A., Gilbert,D.J.,
            Copeland,N.G. and Fletcher,F.A.
  TITLE     Identification and mapping of Casp7, a cysteine protease resembling
            CPP32 beta, interleukin-1 beta converting enzyme, and CED-3
  JOURNAL   Genomics 40 (1), 86-93 (1997)
  MEDLINE   97224489
REFERENCE   2  (bases 1 to 2602)
  AUTHORS   Juan,T.S.C., McNiece,I.K., Argento,J.M., Jenkins,N.A.,
            Gilbert,D.J., Copeland,N.G. and Fletcher,F.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-1996) Developmental Hematology, Amgen, Inc., 1840
            DeHavilland Dr., Thousand Oaks, CA 91320, USA
FEATURES             Location/Qualifiers
     source          1. .2602
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             229. .1239
                     /note="similar to ICE and CED-3 cysteine protease"
                     /codon_start=1
                     /product="Lice2 beta cysteine protease"
                     /protein_id="AAC51152.1"
                     /db_xref="GI:1894913"
                     /translation="MDCVGWPPGRKWHLEKNTSCGGSSGICASYVTQMADDQGCIEEQ
                     GVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVTMRSIKTTRDRVPTYQYNMNFEKLGK
                     CIIINNKNFDKVTGMGVRNGTDKDAEALFKCFRSLGFDVIVYNDCSCAKMQDLLKKAS
                     EEDHTNAACFACILLSHGEENVIYGKDGVTPIKDLTAHFRGDRCKTLLEKPKLFFIQA
                     CRGTELDDGIQADSGPINDTDANPRYKIPVEADFLFAYSTVPGYYSWRSPGRGSWFVQ
                     ALCSILEEHGKDLEIMQILTRVNDRVARHFESQSDDPHFHEKKQIPCVVSMLTKELYF
                     SQ"
BASE COUNT      768 a    515 c    616 g    702 t      1 others
ORIGIN


  Query Match             73.5%;  Score 1007.2;  DB 9;  Length 2602;
  Best Local Similarity   97.7%;  Pred. No. 2.5e-266;
  Matches 1043;  Conservative    0;  Mismatches   23;  Indels    2;  Gaps    2;

Qy    304 GTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGGGCTGTATT 363
          ||  ||||   |   |  || |   ||      ||||||||||| ||||||||||||
Db    293 GTAGCAGTGGGATTTGTGCTTCTTATGTTACCCAGATGGCAGATGA-TCAGGGCTGTATT 351

Qy    364 GAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCGG 423
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    352 GAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCGG 411

Qy    424 TCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATC 483
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    412 TCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATC 471

Qy    484 AAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGC 543
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    472 AAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGC 531

Qy    544 AAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAAC 603
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    532 AAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAAC 591

Qy    604 GGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTG 663
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    592 GGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTG 651

Qy    664 ATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAA 723
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    652 ATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAA 711

Qy    724 GAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAAT 783
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    712 GAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAAT 771

Qy    784 GTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGG 843
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    772 GTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGG 831

Qy    844 GATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGG 903
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    832 GATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGG 891

Qy    904 ACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAAT 963
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    892 ACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAAT 951

Qy    964 CCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGC 1023
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    952 CCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGC 1011

Qy   1024 TATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATC 1083
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1012 TATTACTCATGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATC 1071

Qy   1084 CTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGACAG 1143
          |||||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||
Db   1072 CTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCAGGGTGAATGACAG 1130

Qy   1144 AGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGAT 1203
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1131 AGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGAT 1190

Qy   1204 CCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCAGGG 1263
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1191 CCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCAGGG 1250

Qy   1264 GTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGCTCT 1323
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1251 GTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGCTCT 1310

Qy   1324 TGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1311 TGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1358


RESULT   9
HSU39613
LOCUS       HSU39613                912 bp    mRNA    linear   PRI 19-JAN-1996
DEFINITION  Human cysteine protease ICE-LAP3 mRNA, complete cds.
ACCESSION   U39613
VERSION     U39613.1  GI:1125072
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 912)
  AUTHORS   Duan,H., Chinnaiyan,A.M., Hudson,P.L., Wing,J.P., He,W.W. and
            Dixit,V.M.
  TITLE     ICE-LAP3, a novel mammalian homologue of the Caenorhabditis elegans
```

cell death protein Ced-3 is activated during Fas- and tumor necrosis factor-induced apoptosis
JOURNAL J. Biol. Chem. 271 (3), 1621-1625 (1996)
MEDLINE 96139498
REFERENCE 2 (bases 1 to 912)
AUTHORS Duan,H., Chinnaiyan,A.M., Hudson,P.L., Wing,J.P., He,W.W. and Dixit,V.M.
TITLE Direct Submission
JOURNAL Submitted (27-OCT-1995) Hangjun Duan, Pathology, University of Michigan, 1150 W Medical Center Dr MSRBI Room 7514, Ann Arbor, MI 48109, USA

```
FEATURES             Location/Qualifiers
     source          1. .912
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .912
                     /function="cysteine protease"
                     /note="member of ICE/GED-3 gene family of cysteine
                     proteases"
                     /codon_start=1
                     /product="ICE-LAP3"
                     /protein_id="AAC50346.1"
                     /db_xref="GI:1125073"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENVIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDAIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPHFHEKKQIPCVVSMLTKELYFSQ"
BASE COUNT      263 a    219 c    223 g    207 t
ORIGIN
```

Query Match 64.7%; Score 886.8; DB 9; Length 912;
Best Local Similarity 99.7%; Pred. No. 3e-233;
Matches 899; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
Qy    351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
Db     12 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 71

Qy    411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
Db     72 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 131

Qy    471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
Db    132 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 191

Qy    531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
Db    192 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 251

Qy    591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
Db    252 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 311

Qy    651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
Db    312 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 371

Qy    711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
Db    372 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 431

Qy    771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
Db    432 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 491

Qy    831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
Db    492 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 551

Qy    891 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 950
Db    552 GGCTTGCCGAGGGACCGAGCTTGATGATGCCATCCAGGCCGACTCGGGGCCCATCAATGA 611

Qy    951 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 1010
Db    612 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 671

Qy   1011 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 1070
Db    672 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 731

Qy   1071 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCA 1130
Db    732 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCA 790

Qy   1131 GGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 1190
Db    791 GGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 850

Qy   1191 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAAT 1250
Db    851 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAAT 910

Qy   1251 AG 1252
Db    911 AG 912
```

RESULT 10
HSU37449
LOCUS HSU37449 2378 bp mRNA linear PRI 13-DEC-1995
DEFINITION Human Mch3 isoform beta (Mch3) mRNA, complete cds.
ACCESSION U37449
VERSION U37449.1 GI:1117848
KEYWORDS Mch3 isoform beta.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2378)
AUTHORS Fernandes-Alnemri,T., Litwack,G. and Alnemri,E.S.
TITLE CPP32, a novel human apoptotic protein with homology to Caenorhabditis elegans cell death protein Ced-3 and mammalian interleukin-1 beta-converting enzyme
JOURNAL J. Biol. Chem. 269 (49), 30761-30764 (1994)
MEDLINE 95074098
REFERENCE 2 (bases 1 to 2378)
AUTHORS Fernandes-Alnemri,T., Takahashi,A., Armstrong,R., Krebs,J., Fritz,L., Tomaselli,K.J., Wang,L., Yu,Z., Croce,C.M., Salveson,G., Earnshaw,W.C., Litwack,G. and Alnemri,E.S.
TITLE Mch3, a novel human apoptotic cysteine protease highly related to CPP32
JOURNAL Cancer Res. 55 (24), 6045-6052 (1995)
MEDLINE 96105019
REFERENCE 3 (bases 1 to 2378)
AUTHORS Alnemri,E.S.
TITLE Direct Submission
JOURNAL Submitted (03-OCT-1995) Emad S Alnemri, Pharmacology, Thomas Jefferson University, Jefferson Cancer Institute, 233, S. Tenth Street, Philadelphia, PA 19107, USA

```
FEATURES             Location/Qualifiers
     source          1. .2378
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="T50-6A.1"
                     /cell_line="Jurkat"
                     /cell_type="T-lymphocyte"
     gene            1. .2378
                     /gene="Mch3"
     CDS             147. .908
                     /gene="Mch3"
                     /note="does not have protease activity"
```

```
                     /codon_start=1
                     /evidence=experimental
                     /product="Mch3 isoform beta"
                     /protein_id="AAC50304.1"
                     /db_xref="GI:1117849"
                     /translation="MADDQGCIEEQGVEDSANEDSVDAKPDRSSFVPSLFSKKKKNVT
                     MRSIKTTRDRVPTYQYNMNFEKLGKCIIINNKNFDKVTGMGVRNGTDKDAEALFKCFR
                     SLGFDVIVYNDCSCAKMQDLLKKASEEDHTNAACFACILLSHGEENMESCSVTQAGVQ
                     RRDLGRLQPPPPRLAEGPSLMMASRPTRGPSMTQMLILDTRSQWKLTSSSPIPRFQAI
                     TRGGAQEEAPGLCKPSAPSWRSTEKTWKSCRSSPG"
     polyA_site      2378
                     /gene="Mch3"
BASE COUNT      680 a    485 c    566 g    647 t
ORIGIN


  Query Match           62.6%;  Score 858;  DB 9;  Length 2378;
  Best Local Similarity  92.6%;  Pred. No. 3.3e-225;
  Matches  958;  Conservative    0;  Mismatches   40;  Indels   36;  Gaps    4;

Qy     338 GATGGCAGATGATTCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAG 397
           ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db     146 GATGGCAGATGA-TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAG 204

Qy     398 ATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGA 457
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     205 ATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGA 264

Qy     458 AGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGT 517
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     265 AGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGT 324

Qy     518 ACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATA 577
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     325 ACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATA 384

Qy     578 AAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGT 637
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     385 AAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGT 444

Qy     638 GCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGC 697
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     445 GCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGC 504

Qy     698 AAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCA 757
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     505 AAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCA 564

Qy     758 TCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAA 817
           |||||||||||||||||||||||| |         |||   || | |  |||| |
Db     565 TCCTCTTAAGCCATGGAGAAGAAAATAT--------GGAATCTTGCTCTGTCACCCAGGC 616

Qy     818 AGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAAC 877
           ||| |  |  | |    |  |                    |||  |   |
Db     617 TGGAGTGCAGCGGCGTGATCTCG----------------------GAAGACTGCAA 650

Qy     878 TCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGG 937
           || |  |   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     651 CCTCCACCTCCCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGG 710

Qy     938 GGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCC 997
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     711 GGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCC 770

Qy     998 TCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCT 1057
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     771 TCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCT 830

Qy    1058 GGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGC 1117
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     831 GGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGC 890

Qy    1118 AAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGAC 1177
           | |||||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     891 AGATCCT-CACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGAC 949

Qy    1178 CCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTC 1237
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     950 CCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTC 1009

Qy    1238 TACTTCAGTCAATAGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCAT 1297
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1010 TACTTCAGTCAATAGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCAT 1069

Qy    1298 TAATGAATCACATTTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTT 1357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1070 TAATGAATCACATTTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTT 1129

Qy    1358 CAGGAAAATGTATT 1371
           ||||||||||||||
Db    1130 CAGGAAAATGTATT 1143


RESULT  11
MMCASP7
LOCUS       MMCASP7                 2372 bp    mRNA    linear   ROD 27-MAY-1997
DEFINITION  M.musculus mRNA for caspase-7.
ACCESSION   Y13088
VERSION     Y13088.1  GI:2094813
KEYWORDS    caspase-7; cysteine protease.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2372)
  AUTHORS   Van de Craen,M., Vandenabeele,P., Declercq,W., Van den Brande,I.,
            Van Loo,G., Molemans,F., Schotte,P., Van Criekinge,W., Beyaert,R.
            and Fiers,W.
  TITLE     Characterization of seven murine caspase family members
  JOURNAL   FEBS Lett. 403 (1), 61-69 (1997)
  MEDLINE   97190206
REFERENCE   2  (bases 1 to 2372)
  AUTHORS   Van de Craen,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAY-1997) M. Van de Craen, University of Gent,
            Laboratory of Molecular Biology, K.L. Ledeganckstraat 35, B-9000
            Gent, BELGIUM
FEATURES             Location/Qualifiers
     source          1. .2372
                     /organism="Mus musculus"
                     /strain="C3H/An"
                     /db_xref="taxon:10090"
                     /cell_line="L929r2"
                     /cell_type="fibrosarcoma"
     CDS             190. .1101
                     /standard_name="casp-7"
                     /function="cysteine protease"
                     /codon_start=1
                     /product="caspase-7"
                     /protein_id="CAA73530.1"
                     /db_xref="GI:2094814"
                     /db_xref="SWISS-PROT:P97864"
                     /translation="MTDDQDCAAELEKVDSSSEDGVDAKPDRSSIISSILLKKKRNAS
                     AGPVRTGRDRVPTYLYRMDFQKMGKCIIINNKNFDKATGMDVRNGTDKDAGALFKCFQ
                     NLGFEVTVHNDCSCAKMQDLLRKASEEDHSNSACFACVLLSHGEEDLIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDIDANPRNKIPVEAD
                     FLFAYSTVPGYYSWRNPGKGSWFVQALCSILNEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPRFNEKKQIPCMVSMLTKELYFSR"
BASE COUNT      608 a    588 c    574 g    602 t
ORIGIN


  Query Match           49.5%;  Score 679.2;  DB 10;  Length 2372;
  Best Local Similarity  75.5%;  Pred. No. 6.4e-176;
  Matches  914;  Conservative    0;  Mismatches  278;  Indels   18;  Gaps    5;
```

```
Qy    163 CGCACTACCCCGAGCCAGGGGCGGTGCAAGCCCCGCCCGGCCCTACCCAGGGCGGCTCCT 222
Db     27 CGCTCGCCCCCAGCTTTCCCGGGCTGCAGGCACAGCCCCGTCCGGCCC--GCCCGCTCCT  84

Qy    223 CCCTCCGCAGCGCCGAGACTTTTAGTTTCGCTTTCGCTAAAGGGGCCCCAGACCCTTGCT 282
Db     85 CCCCCTCCGCCTCCGGGACTTTTGCTTTCAGTTTCCCAAAGCTGCCCTCGACCCTTGCG 144

Qy    283 GCGGAGCGACGGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGG 342
Db    145 GAGGACGGACGCTG----------CCGTGGGCTCCTGGCCGCCGCCGTGGGAACGATGA 193

Qy    343 CAGATGATTCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCA 402
Db    194 CCGATGA-TCAGGACTGTGCTGCGGAGCTGGAAAAGGTGGATTCTTCCAGCGAAGACGGA 252

Qy    403 GTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAA 462
Db    253 GTTGACGCCAAGCCAGACCGCTCCTCTATCATCTCCTCTATTCTCTTGAAGAAGAAGAGA 312

Qy    463 AATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAAC 522
Db    313 AATGCCTCTGCGGGCCCCGTCAGGACCGGCCGGGACCGAGTGCCCACTTATCTGTACCGC 372

Qy    523 ATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTG 582
Db    373 ATGGATTTCCAGAAGATGGGTAAATGCATCATCATAAACAACAAGAACTTCGACAAAGCG 432

Qy    583 ACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTC 642
Db    433 ACAGGTATGGACGTCCGGAATGGGACGGACAAAGATGCAGGGGCCCTCTTCAAGTGCTTC 492

Qy    643 CGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGAT 702
Db    493 CAAAACCTGGGTTTTGAAGTAACCGTCCACAATGACTGCTCTTGTGCAAAGATGCAAGAT 552

Qy    703 CTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTC 762
Db    553 CTGCTTAGAAAAGCCTCTGAGGAGGACCACAGCAACTCGGCCTGCTTCGCCTGCGTCCTG 612

Qy    763 TTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGAT 822
Db    613 CTGAGCCACGGGGAAGAGGACCTGATTTACGGGAAAGATGGCGTGACACCCATAAAGGAT 672

Qy    823 TTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTC 882
Db    673 CTGACAGCTCATTTTAGGGGAGACCGATGCAAAACCCTGTTAGAGAAACCCAAACTCTTC 732

Qy    883 TTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCC 942
Db    733 TTCATTCAGGCATGCCGAGGGACGGAGCTCGATGATGGAATCCAGGCTGACTCGGGGCCC 792

Qy    943 ATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTC 1002
Db    793 ATCAACGACATTGACGCTAATCCCCGCAACAAGATCCCGGTGGAAGCCGACTTCCTCTTT 852

Qy   1003 GCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTT 1062
Db    853 GCTTACTCCACGGTTCCAGGTTATTACTCATGGAGGAACCCAGGGAAAGGCTCCTGGTTT 912

Qy   1063 GTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATC 1122
Db    913 GTGCAGGCCCTCTGCTCCATCCTGAATGAGCATGGCAAGGACCTCGAGATCATGCAGATC 972

Qy   1123 CTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACA 1182
Db    973 CTG-ACCAGGGTGAACGACAGGGTGGCCAGGCACTTCGAGTCCCAGTCTGATGATCCACG 1031

Qy   1183 CTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTT 1242
Db   1032 CTTCAACGAGAAGAAGCAGATCCCGTGTATGGTGTCCATGCTCACCAAAGAGCTGTACTT 1091

Qy   1243 CAGTCAATAGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATG 1302
```

```
Db   1092 CAGCCGTTGACCACCCTTCAGCTGAGAACCTGCCGCCGTTCGTTGATGAATCCAGTTTTT 1151

Qy   1303 AATCACATTTTTTT---ATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCA 1359
Db   1152 ATTTTATTTTTGTTCCGATGCTCTCAAAATATCCAGAAATGTTGAGGGGATTTAATTTCA 1211

Qy   1360 GGAAAATGTA 1369
Db   1212 GGAAAGTCTA 1221


RESULT  12
AF072124
LOCUS       AF072124                1063 bp    mRNA    linear   ROD 24-MAY-2000
DEFINITION  Rattus norvegicus caspase-7 mRNA, complete cds.
ACCESSION   AF072124
VERSION     AF072124.1  GI:3243275
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1063)
  AUTHORS   Forghani,F. and Roy,S.
  TITLE     Rat caspase-7 sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1063)
  AUTHORS   Forghani,F. and Roy,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-1998) Biochemistry, Merck Frosst Centre for
            Therapeutic Research, P.O. Box 1005, Pointe-Claire-Dorval, QUE H9H
            3L1, Canada
FEATURES             Location/Qualifiers
     source          1. .1063
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /tissue_type="spleen"
     CDS             3. .914
                     /codon_start=1
                     /product="caspase-7"
                     /protein_id="AAC24011.1"
                     /db_xref="GI:3243276"
                     /translation="MTDDQDCAAELEMADSSTEDGVDAKPDRSTIISSLLWKKKKNAS
                     MCPVSTTRDRVPTYLYRMDFEKMGKCIIINNKNFDKATGMDVRNGTDKDAEALFKCFR
                     SLGFEVTVYNDCSCAKMQDLLRRASEEDHSNSACFACVLLSHGEENLIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDTDANPRYKIPVEAD
                     FLFAYSTVPGYYSWRNPGKGSWFVQALCSILNEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPRFNEKKQIPCMVSMLTKELYFGR"
BASE COUNT      278 a    268 c    275 g    242 t
ORIGIN


  Query Match             48.1%;  Score 660;  DB 10;  Length 1063;
  Best Local Similarity   79.1%;  Pred. No. 1e-170;
  Matches  810;  Conservative    0;  Mismatches  210;  Indels    4;  Gaps    2;

Qy    351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
Db     14 TCAGGACTGTGCTGCAGAGCTGGAAATGGCAGATTCTTCGACGGAAGACGGAGTTGACGC  73

Qy    411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
Db     74 TAAGCCAGACCGCTCCACCATCATCTCATCCCTTCTCTGGAAGAAGAAGAAAAATGCCTC 133

Qy    471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
Db    134 TATGTGCCCCGTCAGTACCACCCGGGACCGCGTGCCCACTTACCTGTACCGTATGGACTT 193

Qy    531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
Db    194 CGAGAAGATGGGCAAATGTATCATCATAAATAACAAGAACTTCGACAAAGCGACAGGTAT 253
```

```
Qy     591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
Db     254 GGACGTCCGGAATGGGACGGACAAAGACGCAGAGGCCCTCTTCAAGTGCTTCCGAAGCCT 313

Qy     651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
Db     314 GGGTTTTGAAGTGACTGTCTATAATGACTGCTCTTGCGCAAAGATGCAGGATCTGCTTAG 373

Qy     711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
Db     374 ACGAGCCTCTGAAGAGGACCACAGCAACTCAGCCTGCTTCGCCTGCGTCCTGCTGAGCCA 433

Qy     771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
Db     434 CGGAGAAGAGAATCTGATTTATGGGAAAGACGGCGTGACACCGATAAAGGATCTGACAGC 493

Qy     831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
Db     494 TCATTTTAGGGGAGACCGATGCAAAACCCTGCTAGAGAAACCCAAGCTCTTCTTCATCCA 553

Qy     891 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 950
Db     554 GGCGTGCCGAGGGACGGAGCTGGATGACGGGATCCAGGCTGACTCGGGGCCTATCAACGA 613

Qy     951 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 1010
Db     614 CACCGACGCTAATCCCCGCTACAAGATCCCGGTGGAAGCTGACTTTCTCTTTGCTTACTC 673

Qy    1011 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 1070
Db     674 CACGGTTCCAGGCTATTATTCGTGGAGGAACCCAGGAAAGGGCTCCTGGTTTGTGCAGGC 733

Qy    1071 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCA 1130
Db     734 CCTCTGCTCCATCCTGAATGAGCATGGCAAGGACCTGGAGATCATGCAGATCCT-GACCA 792

Qy    1131 GGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 1190
Db     793 GGGTGAACGACAGGGTGGCCAGACACTTCGAGTCCCAGTCTGATGACCCCCGCTTCAATG 852

Qy    1191 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAAT 1250
Db     853 AGAAGAAGCAGATTCCATGTATGGTGTCCATGCTCACCAAAGAGCTCTACTTTGGCCGTT 912

Qy    1251 AGCCATATCAGGGGTACATTCTAGCTGAGAAGCA---ATGGGTCACTCATTAATGAATCA 1307
Db     913 GACCAGCCTCCAACTGAGAACCTGCTGTCATTCACTGATGAATCCAGCTTTTATTTTATT 972

Qy    1308 CATTTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATG 1367
Db     973 TTTGTTTTGATGCTCTCAAAATATCCAGAAATGTTAAGGAGATTTAATTTCAGGAAAGTA 1032

Qy    1368 TATT 1371
Db    1033 TATT 1036
```

```
RESULT  13
BC005428
LOCUS       BC005428               2305 bp    mRNA    linear   ROD 12-JUL-2001
DEFINITION  Mus musculus, caspase 7, clone MGC:5872 IMAGE:3600155, mRNA,
            complete cds.
ACCESSION   BC005428
VERSION     BC005428.1  GI:13529367
KEYWORDS    MGC.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2305)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 11 Row: b Column: 10
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 2094813.
FEATURES             Location/Qualifiers
     source          1. .2305
                     /organism="Mus musculus"
                     /db_xref="LocusID:12369"
                     /db_xref="taxon:10090"
                     /clone="MGC:5872 IMAGE:3600155"
                     /tissue_type="Mammary tumor. Metallothionien-TGF alpha
                     model. 10 month old virgin mouse. Taken by biopsy."
                     /clone_lib="NCI_CGAP_Mam1"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             47. .958
                     /codon_start=1
                     /product="caspase 7"
                     /protein_id="AAH05428.1"
                     /db_xref="GI:13529368"
                     /translation="MTDDQDCAAELEKVDSSSEDGVDAKPDRSSIISSILLKKKRNAS
                     AGPVRTGRDRVPTYLYRMDFQKMGKCIIINNKNFDKATGMDVRNGTDKDAGALFKCFQ
                     NLGFEVTVHNDCSCAKMQDLLRKASEEDHSNSACFACVLLSHGEEDLIYGKDGVTPIK
                     DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGIQADSGPINDIDANPRNKIPVEAD
                     FLFAYSTVPGYYSWRNPGKGSWFVQALCSILNEHGKDLEIMQILTRVNDRVARHFESQ
                     SDDPRFNEKKQIPCMVSMLTKELYFSR"
BASE COUNT      665 a    523 c    538 g    579 t
ORIGIN
```

```
  Query Match           47.8%;  Score 656;  DB 10;  Length 2305;
  Best Local Similarity   77.3%;  Pred. No. 1.6e-169;
  Matches  835;  Conservative    0;  Mismatches  240;  Indels    5;  Gaps    3;

Qy     293 GGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTC 352
Db       1 GGAGGACGGACGCTGCCGTGGGCTCCTGGCCGCCGCCGTGGGAACGATGACCGATGA-TC 59

Qy     353 AGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTA 412
Db      60 AGGACTGTGCTGCGGAGCTGGAAAAGGTGGATTCTTCCAGCGAAGACGGAGTTGACGCCA 119

Qy     413 AGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCA 472
Db     120 AGCCAGACCGCTCCTCTATCATCTCCTCTATTCTCTTGAAGAAGAAGAGAAATGCCTCTG 179

Qy     473 TGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTG 532
Db     180 CGGGCCCCGTCAGGACCGGCCGGGACCGAGTGCCCACTTATCTGTACCGCATGGATTTCC 239

Qy     533 AAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGG 592
Db     240 AGAAGATGGGTAAATGCATCATCATAAACAACAAGAACTTCGACAAAGCGACAGGTATGG 299

Qy     593 GCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGG 652
```

```
Db    300 ACGTCCGGAATGGGACGGACAAAGATGCAGGGGCCCTCTTCAAGTGCTTCCAAAACCTGG 359

Qy    653 GTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAA 712

Db    360 GTTTTGAAGTAACCGTCCACAATGACTGCTCTTGTGCAAAGATGCAAGATCTGCTTAGAA 419

Qy    713 AAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATG 772

Db    420 AAGCCTCTGAGGAGGACCACAGCAACTCGGCCTGCTTCGCCTGCGTCCTGCTGAGCCACG 479

Qy    773 GAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCC 832

Db    480 GGGAAGAGGACCTGATTTACGGGAAAGATGGCGTGACACCCATAAAGGATCTGACAGCTC 539

Qy    833 ACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGG 892

Db    540 ATTTTAGGGGAGACCGATGCAAAACCCTGTTAGAGAAACCCAAACTCTTCTTCATTCAGG 599

Qy    893 CTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACA 952

Db    600 CATGCCGAGGGACGGAGCTCGATGATGGAATCCAGGCTGACTCGGGGCCCATCAACGACA 659

Qy    953 CAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCA 1012

Db    660 TTGACGCTAATCCCCGCAACAAGATCCCGGTGGAAGCCGACTTCCTCTTTGCTTACTCCA 719

Qy   1013 CGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCC 1072

Db    720 CGGTTCCAGGTTATTACTCATGGAGGAACCCAGGGAAAGGCTCCTGGTTTGTGCAGGCCC 779

Qy   1073 TCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGG 1132

Db    780 TCTGCTCCATCCTGAATGAGCATGGCAAGGACCTCGAGATCATGCAGATCCT-GACCAGG 838

Qy   1133 GTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAG 1192

Db    839 GTGAACGACAGGGTGGCCAGGCACTTCGAGTCCCAGTCTGATGATCCACGCTTCAACGAG 898

Qy   1193 AAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAG 1252

Db    899 AAGAAGCAGATCCCGTGTATGGTGTCCATGCTCACCAAAGAGCTGTACTTCAGCCGTTGA 958

Qy   1253 CCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTT 1312

Db    959 CCACCCTTCAGCTGAGAACCTGCCGCCGTTCGTTGATGAATCCAGTTTTTATTTTATTTT 1018

Qy   1313 TTTT---ATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTA 1369

Db   1019 TGTTCCGATGCTCTCAAAATATCCAGAAATGTTGAGGGGATTTAATTTCAGGAAAGTCTA 1078


RESULT  14
MMU67321
LOCUS       MMU67321                2006 bp    mRNA    linear   ROD 19-MAR-1997
DEFINITION  Mus musculus Lice2 cysteine protease (Lice2) mRNA, complete cds.
ACCESSION   U67321
VERSION     U67321.1  GI:1894916
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2006)
  AUTHORS   Juan,T.S., McNiece,I.K., Argento,J.M., Jenkins,N.A., Gilbert,D.J.,
            Copeland,N.G. and Fletcher,F.A.
  TITLE     Identification and mapping of Casp7, a cysteine protease resembling
            CPP32 beta, interleukin-1 beta converting enzyme, and CED-3
  JOURNAL   Genomics 40 (1), 86-93 (1997)
  MEDLINE   97224489
REFERENCE   2  (bases 1 to 2006)
  AUTHORS   Juan,T.S.C., McNiece,I.K., Argento,J.M., Jenkins,N.A.,
            Gilbert,D.J., Copeland,N.G. and Fletcher,F.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-1996) Developmental Hematology, Amgen, Inc., 1840
            DeHavilland Dr., Thousand Oaks, CA 91320, USA
FEATURES             Location/Qualifiers
     source          1. .2006
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="19"
                     /map="Between Adra2a and Aop1"
     gene            1. .2006
                     /gene="Lice2"
     CDS             474. .1496
                     /gene="Lice2"
                     /note="similar to ICE and CED-3 cysteine protease"
                     /codon_start=1
                     /product="Lice2 cysteine protease"
                     /protein_id="AAC53068.1"
                     /db_xref="GI:1894917"
                     /translation="MFTAQVFSESFTKTELLPSTLAEDGRCRGLLAAAVGTMTDDQDC
                     AAELEKVDSSSEDGVDAKPDRSSIISSILLKKKRNASAGPVRTGRDRVPTYLYRMDFQ
                     KMGKCIIINNKNFDKATGMDVRNGTDKDAGALFKCFQNLGFEVTVHNDCSCAKMQDLL
                     RKASEEDHSNSACFACVLLSHGEEDLIYGKDGVTPIKDLTAHFRGDRCKTLLEKPKLF
                     FIQACRGTELDDGIQADSGPINDIDANPRNKIPVEADFLFAYSTVPGYYSWRNPGKGS
                     WFVQALCSILNEHGKDLEIMQILTRVNDRVARHFESQSDDPRFNEKKQIPCMVSMLTK
                     ELYFSR"
BASE COUNT      543 a    490 c    462 g    511 t
ORIGIN


  Query Match           47.8%;  Score 655.4;  DB 10;  Length 2006;
  Best Local Similarity  77.2%;  Pred. No. 2.2e-169;
  Matches  835;  Conservative    0;  Mismatches  241;  Indels    5;  Gaps    3;

Qy     292 CGGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATT 351

Db     538 CGGAGGACGGACGCTGCCGTGGGCTCCTGGCCGCCGCCGTGGGAACGATGACCGATGA-T 596

Qy     352 CAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCT 411

Db     597 CAGGACTGTGCTGCGGAGCTGGAAAAGGTGGATTCTTCCAGCGAAGACGGAGTTGACGCC 656

Qy     412 AAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACC 471

Db     657 AAGCCAGACCGCTCCTCTATCATCTCCTCTATTCTCTTGAAGAAGAAGAGAAATGCCTCT 716

Qy     472 ATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTT 531

Db     717 GCGGGCCCCGTCAGGACCGGCCGGGACCGAGTGCCCACTTATCTGTACCGCATGGATTTC 776

Qy     532 GAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATG 591

Db     777 CAGAAGATGGGTAAATGCATCATCATAAACAACAAGAACTTCGACAAAGCGACAGGTATG 836

Qy     592 GGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTG 651

Db     837 GACGTCCGGAATGGGACGGACAAAGATGCAGGGGCCCTCTTCAAGTGCTTCCAAAACCTG 896

Qy     652 GGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAA 711

Db     897 GGTTTTGAAGTAACCGTCCACAATGACTGCTCTTGTGCAAAGATGCAAGATCTGCTTAGA 956

Qy     712 AAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCAT 771

Db     957 AAAGCCTCTGAGGAGGACCACAGCAACTCGGCCTGCTTCGCCTGCGTCCTGCTGAGCCAC 1016

Qy     772 GGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCC 831

Db    1017 GGGGAAGAGGACCTGATTTACGGGAAAGATGGCGTGACACCCATAAAGGATCTGACAGCT 1076

Qy     832 CACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAG 891

Db    1077 CATTTTAGGGGAGACCGATGCAAAACCCTGTTAGAGAAACCCAAACTCTTCTTCATTCAG 1136
```

```
Qy     892 GCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGAC 951
           || |||||||||||| |||||  || ||||| |||||||| ||||||||||||||||||| |||
Db    1137 GCATGCCGAGGGACGGAGCTCGACGATGGAATCCAGGCTGACTCGGGGCCCATCAACGAC 1196

Qy     952 ACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCC 1011
           |  || ||||||||  ||  |||||||||| |||||||| |||||||||| || || |||
Db    1197 ATTGACGCTAATCCCCGCAACAAGATCCCGGTGGAAGCCGACTTCCTCTTTGCTTACTCC 1256

Qy    1012 ACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCC 1071
           |||||||||||  |||||||| |||||||  |||||| | |||||||||||||||||| |||
Db    1257 ACGGTTCCAGGTTATTACTCATGGAGGAACCCAGGGAAAGGCTCCTGGTTTGTGCAGGCC 1316

Qy    1072 CTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAG 1131
           |||||||||||||||  | |||||  || || |||||  || ||||||||  |||||  |||||
Db    1317 CTCTGCTCCATCCTGAATGAGCATGGCAAGGACCTCGAGATCATGCAGATCCT-GACCAG 1375

Qy    1132 GGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGA 1191
           ||||||  ||||  || |||||||||| |||||  |||||||||||  ||||  ||||  | ||
Db    1376 GGTGAACGACAGGGTGGCCAGGCACTTCGAGTCCCAGTCTGATGATCCACGCTTCAACGA 1435

Qy    1192 GAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATA 1251
           ||||||||||||||| ||| |||| |||||||||||||| || || |||||||  |   |
Db    1436 GAAGAAGCAGATCCCGTGTATGGTGTCCATGCTCACCAAAGAGCTGTACTTCAGCCGTTG 1495

Qy    1252 GCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATT 1311
            |||      | |     |  || |     |  ||    |  || |  | | |    ||
Db    1496 ACCACCCTTCAGCTGAGAACCTGCCGCCGTTCGTTGATGAATCCAGTTTTTATTTTATTT 1555

Qy    1312 TTTTT---ATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGT 1368
           || ||   |||||||  |||||| |||||||  |   ||  |||||||||||||| | |
Db    1556 TTGTTCCGATGCTCTCAAAATATCCAGAAATGTTGAGGGGATTTAATTTCAGGAAAGTCT 1615

Qy    1369 A 1369
           |
Db    1616 A 1616
```

```
RESULT  15
MAU47332
LOCUS       MAU47332                1939 bp    mRNA    linear   ROD 05-JUN-1996
DEFINITION  Mesocricetus auratus MCH3/SCA-2 mRNA, complete cds.
ACCESSION   U47332
VERSION     U47332.1  GI:1245143
KEYWORDS    .
SOURCE      golden hamster.
  ORGANISM  Mesocricetus auratus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Mesocricetus.
REFERENCE   1  (bases 1 to 1939)
  AUTHORS   Pai,J.T., Brown,M.S. and Goldstein,J.L.
  TITLE     Purification and cDNA cloning of a second apoptosis-related
            cysteine protease that cleaves and activates sterol regulatory
            element binding proteins
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 93 (11), 5437-5442 (1996)
  MEDLINE   96224303
REFERENCE   2  (bases 1 to 1939)
  AUTHORS   Pai,J.-T., Brown,M.S. and Goldstein,J.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-1996) Jih-tung Pai, Molecular Genetics,
            University of Texas, Southwestern Medical Center, 5323 Harry Hines
            Blvd., Dallas, TX 75235-9046, USA
FEATURES             Location/Qualifiers
     source          1. .1939
                     /organism="Mesocricetus auratus"
                     /db_xref="taxon:10036"
     CDS             59. .970
                     /note="cysteine protease"
                     /codon_start=1
                     /product="MCH3/SCA-2"
                     /protein_id="AAC52595.1"
                     /db_xref="GI:1245144"
                     /translation="MADDQNCAPELEKADPSGEDGVDAKPDRSSIISSILGKKKKNAS
                                   ACPVKTARDRVPTYLYRMDFEKMGKCIIINNKNFDKVTGMDVRNGTDKDAEALFKCFR
                                   SLGFDVVVYNDCSCAKMQDLLRKASEEDHSNSACFACVLLSHGEENLIYGKDGVTPIK
                                   DLTAHFRGDRCKTLLEKPKLFFIQACRGTELDDGVQADSGPINETDANPRYKIPVEAD
                                   FLFAYSTVPGYYSWRNPGKGSWFVQALCSILDEHGKDLEIMQILTRVNDRVARHFESQ
                                   CDDPCFNEKKQIPCMVSMLTKELYFGR"
BASE COUNT      493 a    464 c    523 g    459 t
ORIGIN
```

```
  Query Match           46.2%;  Score 632.8;  DB 10;  Length 1939;
  Best Local Similarity  77.3%;  Pred. No. 3.8e-163;
  Matches  829;  Conservative    0;  Mismatches  217;  Indels   26;  Gaps    4;

Qy     318 CGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGGGCTGTATTGAAGAGCAGGGGGT 377
           | | | ||||  | |  ||||||||||| || ||| |||| | ||||| || |
Db      38 CCCGCCACCGTGGGGACGACGATGGCAGACGA-CCAGAACTGTGCTCCAGAGCTGGAGAA 96

Qy     378 TGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACC 437
           | |||| |  |  ||||||    || |||||||||| ||||| |||||  || |  |
Db      97 AGCGGATCCTTCCGGTGAAGACGGGGTCGATGCTAAGCCCGACCGCTCCTCCATTATCTC 156

Qy     438 GTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGA 497
           ||  |  || | |||||||||||||||| | |  | |  || |||||||  |||||||
Db     157 ATCGATTCTCGGGAAGAAGAAGAAAAATGCCTCTGCGTGCCCCGTCAAGACTGCCCGGGA 216

Qy     498 CCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAAT 557
           |||||||| || || | ||| |  ||| |||| || ||| |||| ||||||||||| ||
Db     217 CCGAGTGCCCACCTACCTGTATAGAATGGATTTCGAGAAGATGGGGAAATGCATCATTAT 276

Qy     558 AAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGA 617
           |||||||||| |||| || |||||||||||||||||| ||| || ||||| ||||||||||
Db     277 AAACAACAAAAACTTCGACAAAGTGACAGGTATGGACGTCCGGAACGGGACAGACAAAGA 336

Qy     618 TGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGA 677
           || ||||| |||||||||||||||| || ||||| |||||||| ||||||| |||||
Db     337 CGCAGAGGCCCTCTTCAAGTGCTTCCGCAGTCTGGGCTTTGACGTGGTTGTCTACAATGA 396

Qy     678 CTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAA 737
           |||||||||||| |||||||||| |||||| |||||| || || |||||||| |  ||
Db     397 CTGCTCTTGTGCGAAGATGCAAGACCTGCTTAGAAAAGCCTCGGAGGAGGACCACAGCAA 456

Qy     738 TGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAA 797
           | |||||||||||||||| |||| | |||| |||||||| ||| |||||| ||||
Db     457 CTCAGCCTGCTTCGCCTGCGTCCTGCTGAGCCACGGAGAAGAGAATCTAATTTACGGGAA 516

Qy     798 AGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAAC 857
           |||||| || || || ||||||||  |||| || || |||||||| ||| | ||||||||
Db     517 AGATGGCGTGACGCCGATAAAGGACCTGACGGCTCATTTTAGGGGAGATCGGTGCAAAAC 576

Qy     858 CCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGA 917
           ||| ||||||||||||||||||||||||||||| |||||  |||||||| || |||||  |||||
Db     577 CCTGTTAGAGAAACCCAAACTCTTCTTCATCCAGGCATGCCGAGGCACGGAGCTCGATGA 636

Qy     918 TGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGAT 977
           ||  |||||||||||||| |||||  |||| || || || || || || || || ||||||||
Db     637 CGGGGTCCAGGCCGACTCTGGGCCTATCAACGAAACTGACGCCAACCCCCGGTACAAGAT 696

Qy     978 CCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAG 1037
           ||| |||||||| |||||||||||||||| || |||||  |||||||||||||||||  |||||
Db     697 CCCGGTGGAAGCCGACTTCCTCTTCGCTTACTCCACAGTTCCAGGCTATTACTCATGGAG 756

Qy    1038 GAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGG 1097
           || ||| |||| ||||||||||||||||||| |||||||| ||||||||||  |||||  ||
Db     757 GAACCCGGGAAAAGGCTCCTGGTTTGTGCAGGCCCTCTGTTCCATCCTGGATGAGCATGG 816

Qy    1098 AAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACT 1157
           ||||||||||| |||||||| |||||  |||| |||| |||| |||| || ||||||||||
Db     817 CAAAGACCTGGAGATCATGCAGATCCT-GACCAGAGTGAACGACAGGGTAGCCAGGCACT 875

Qy    1158 TTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCT 1217
           | |||||  |||| |||  |||||   |||| | ||||||||||||||||| ||| |||| |
Db     876 TCGAGTCCCAGTGTGACGACCCCCTGCTTCAACGAGAAGAAGCAGATCCCGTGTATGGTTT 935
```

```
Qy   1218 CCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCAGGGGTACATTCTAGCTG 1277
          ||||||||||||| || |||||||| | |  |  ||||       | ||  | ||
Db    936 CCATGCTCACCAAAGAGCTCTACTTCGGCCGTTGACCATCCTTCCCTTCCAGCCGAG--- 992

Qy   1278 AGAAGCAATGGGTCACTCATTAATGAATCACATT----------------TTTTTATG 1319
            |||     |||||||||  |||| ||  |||                 |||| |||
Db    993 ---AGCCCGCAGTCACTCATCCATGAGTCCAATTGTTAGTCTGCTTTTATTGTTTGATG 1049

Qy   1320 CTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
          ||||  || ||  | |||||  |  ||   ||||||||||||||| | ||||
Db   1050 CTCTCAAAGTACCCGGAAATGTTAAAGAGATTTAATTTCAGGAAAGTATATT 1101
```

Search completed: August 15, 2002, 23:58:04
Job time: 6105 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2002, 22:21:34 ; Search time 323.26 Seconds
(without alignments)
7281.723 Million cell updates/sec

Title: US-09-613-508-1
Perfect score: 1371
Sequence: 1 GCACGAGAAACTTTGCTGTG.....TAATTTCAGGAAAATGTATT 1371

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1371	100.0	1371	17	AAT15276	Human interleukin-
2	1052.8	76.8	2341	18	AAT66970	Cysteine protease
3	1050.4	76.6	2309	18	AAT66992	Apototic protease
4	1050.4	76.6	2309	22	AAD03914	Human caspase-7 DN
5	888.4	64.8	913	20	AAX81223	Nucleotide sequenc
6	858	62.6	2377	18	AAT66993	Cysteine protease
7	232.8	17.0	1297	22	AAS10510	Mouse caspase 3 DN
8	230.6	16.8	2538	21	AAC64043	pYNGCas3p10p35 ins
9	230.4	16.8	834	22	AAI66510	Pig caspase coding
10	229	16.7	1158	17	AAT33567	Pro-Yama cDNA. Ho
11	229	16.7	1158	19	AAV05471	Nucleic acid encod
12	229	16.7	1159	17	AAT15277	Human interleukin-
13	229	16.7	2635	22	AAS10503	Human caspase 3 DN
14	229	16.7	2635	22	AAD03910	Human caspase-3 cD
15	227.8	16.6	834	22	AAD13075	Mammalia caspase-3
16	227.8	16.6	835	20	AAX81219	Nucleotide sequenc
17	226.4	16.5	6436	22	AAD13078	pcDNA3-1XUb-C3 con
18	226.4	16.5	6703	22	AAD13079	pcDNA3-2XUb-C3 con
19	226.4	16.5	6946	22	AAD13080	pcDNA3-3XUb-C3 con
20	226.4	16.5	7189	22	AAD13081	pcDNA3-4XUb-C3 con
21	226.4	16.5	7247	22	AAD13084	pcDNA3-Ub-Met-C3 c
22	225.8	16.5	1001	19	AAV32608	Mutant human apopa
23	225.8	16.5	1001	19	AAV32615	Mutant human apopa
24	225.2	16.4	1177	19	AAV13959	Rat interleukin-1
25	225.2	16.4	1177	19	AAV13960	Rat interleukin-1
26	212.2	15.5	786	22	AAD03918	Caspase chimera ca
27	139	10.1	2658	20	AAV80061	Aspartate-specific
28	133.2	9.7	858	20	AAX81215	Nucleotide sequenc
29	133.2	9.7	873	20	AAX81214	Nucleotide sequenc
30	131.4	9.6	3614	22	AAH26878	Anti-HER2 Fab reve
31	110	8.0	3008	23	ABL09303	Drosophila melanog
32	107.2	7.8	883	20	AAX81222	Nucleotide sequenc
33	106.2	7.7	1265	23	ABL03125	Drosophila melanog
34	106.2	7.7	3265	23	ABL03124	Drosophila melanog
35	105.6	7.7	1183	21	AAZ98834	Interleukin 1 conv
36	105.6	7.7	1204	21	AAZ98839	Interleukin 1 conv
37	105.6	7.7	1545	18	AAT43537	Apoptotic cysteine
38	105.6	7.7	1545	22	AAD03913	Human caspase-6 DN
39	97.6	7.1	1313	18	AAT43538	Mch2 beta isoform
40	85.8	6.3	1489	22	AAD12324	Mouse caspase 8 DN
41	82.4	6.0	903	20	AAX81216	Nucleotide sequenc
42	82.2	6.0	2619	18	AAT61406	MACH isoform alpha
43	82.2	6.0	2887	18	AAT61405	MACH isoform alpha
44	82.2	6.0	2887	22	AAD03915	Human caspase-8 DN
45	80.6	5.9	870	21	AAZ56946	Human procaspase 8

ALIGNMENTS

RESULT 1
AAT15276
ID AAT15276 standard; cDNA; 1371 BP.
XX
AC AAT15276;
XX
DT 27-OCT-1996 (first entry)
XX
DE Human interleukin-I-converting enzyme-like apoptosis protease-3 cDNA.
XX
KW ICE-LAP-3; interleukin-I-converting enzyme-like apoptosis protease;
KW enzyme; Alzheimer's disease; Parkinson's disease; septic shock;
KW head injury; rheumatoid arthritis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 115..1137
FT /*tag= a
XX
PN WO9613603-A1.
XX
PD 09-MAY-1996.
XX
PF 06-JUN-1995; 95WO-US07235.
XX
PR 01-NOV-1994; 94US-0334251.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Hastings GA, He W, Hudson PL, Rosen CA;
XX

```
DR   WPI; 1996-239509/24.
DR   P-PSDB; AAR95830.
XX
PT   Human ICE-LAP-3 and -4 DNA and protein - useful in the diagnosis
PT   and treatment of Alzheimer's disease, Parkinson's disease,
PT   rheumatoid arthritis, septic shock and head injury
XX
PS   Claim 1; Page 91-92; 67pp; English.
XX
CC   This sequence was derived from a cDNA library from human endometrial
CC   tumuor and may be used for the recombinant production of the ICE-LAP-
CC   3 protein e.g. in COS cells. The encoded protein may be used
CC   therapeutically, e.g. as an antitumor or antiviral agent and to
CC   control embryonic development and tissue homeostasis. The protein
CC   can also be used to treat immunosuppression disorders, such as AIDS,
CC   by targeting virus infected cells for cell death. The DNA may find
CC   use in gene therapy applications.
XX
SQ   Sequence 1371 BP; 355 A; 357 C; 354 G; 305 T; 0 other;
```

```
  Query Match            100.0%;  Score 1371;  DB 17;  Length 1371;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1371;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCACGAGAAACTTTGCTGTGCGCGTTCTCCCGCGCGCGGGCTCAACTTTGTAGAGCGAGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 gcacgagaaactttgctgtgcgcgttctcccgcgcgcgggctcaactttgtagagcgagg 60

Qy      61 GGCCAACTTGGCAGAGCGCGCGGCCAGCTTTGCAGAGAGCGCCCTCCAGGGACTATGCGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ggccaacttggcagagcgcgcggccagctttgcagagagcgccctccagggactatgcgt 120

Qy     121 GCGGGGACACGGGTCGCTTTGGGCTCTTCCACCCCTGCGGAGCGCACTACCCCGAGCCAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gcggggacacgggtcgctttgggctcttccaccccctgcggagcgcactaccccgagccag 180

Qy     181 GGGCGGTGCAAGCCCCGCCCGGCCCTACCCAGGGCGGCTCCTCCCTCCGCAGCGCCGAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 gggcggtgcaagccccgcccggccctacccagggcggctcctccctccgcagcgccgaga 240

Qy     241 CTTTTAGTTTCGCTTTCGCTAAAGGGGCCCCAGACCCTTGCTGCGGAGCGACGGAGAGAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 cttttagtttcgctttcgctaaaggggccccagacccttgctgcggagcgacggagagag 300

Qy     301 ACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGGGCTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 actgtgccagtcccagccgccctaccgccgtgggaacgatggcagatgattcagggctgt 360

Qy     361 ATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 attgaagagcaggggttgaggattcagcaaatgaagattcagtggatgctaagccagac 420

Qy     421 CGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 cggtcctcgtttgtaccgtccctcttcagtaagaagaagaaaaatgtcaccatgcgatcc 480

Qy     481 ATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 atcaagaccacccgggaccgagtgcctacatatcagtacaacatgaattttgaaaagctg 540

Qy     541 GGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 ggcaaatgcatcataataaacaacaagaactttgataaagtgacaggtatgggcgttcga 600

Qy     601 AACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 aacggaacagacaaagatgccgaggcgctcttcaagtgcttccgaagcctgggttttgac 660

Qy     661 GTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 gtgattgtctataatgactgctcttgtgccaagatgcaagatctgcttaaaaaagcttct 720

Qy     721 GAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 gaagaggaccatacaaatgccgcctgcttcgcctgcatcctcttaagccatggagaagaa 780

Qy     781 AATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 aatgtaatttatgggaaagatggtgtcacaccaataaaggatttgacagcccactttagg 840

Qy     841 GGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 ggggatagatgcaaaacccttttagagaaacccaaactcttcttcattcaggcttgccga 900

Qy     901 GGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 gggaccgagcttgatgatggcatccaggccgactcggggcccatcaatgacacagatgct 960

Qy     961 AATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 aatcctcgatacaagatcccagtggaagctgacttcctcttcgcctattccacggttcca 1020

Qy    1021 GGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 ggctattactcgtggaggagcccaggaagaggctcctggtttgtgcaagccctctgctcc 1080

Qy    1081 ATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 atcctggaggagcacggaaaagacctggaaatcatgcaaatcctccaccagggtgaatga 1140

Qy    1141 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 cagagttgccaggcactttgagtctcagtctgatgacccacacttccatgagaagaagca 1200

Qy    1201 GATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 gatccctgtgtggtctccatgctcaccaaggaactctacttcagtcaatagccatatca 1260

Qy    1261 GGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 ggggtacattctagctgagaagcaatgggtcactcattaatgaatcacatttttttatgc 1320

Qy    1321 TCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 tcttgaaatattcagaaattctccaggattttaatttcaggaaaatgtatt 1371
```

```
RESULT   2
AAT66970
ID   AAT66970 standard; cDNA; 2341 BP.
XX
AC   AAT66970;
XX
DT   21-JUL-1997  (first entry)
XX
DE   Cysteine protease CMH-1 cDNA.
XX
KW   Cysteine protease; CMH-1; CPP32/Mch2 homologue-1; apoptosis;
KW   programmed cell death; cancer; neurodegenerative disease;
KW   autoimmune disease; gene therapy; diagnosis; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             54..965
FT                   /*tag= a
FT                   /product= CMH-1 proenzyme
FT                   /note= "bases 123-962 and 138-962 encode active
FT                         active CMH-1 (Claim 2)"
XX
PN   WO9716552-A1.
XX
PD   09-MAY-1997.
XX
PF   01-NOV-1996;   96WO-US17431.
```

XX
PR 16-NOV-1995; 95US-0558733.
PR 03-NOV-1995; 95US-0007211.
PR 06-NOV-1995; 95US-0007251.
XX
PA (VERT-) VERTEX PHARM INC.
XX
PI Lippke JA, Su M;
XX
DR WPI; 1997-272121/24.
DR P-PSDB; AAW15247.
XX
PT DNA encoding active, activatable or inactive cysteine protease CMH-1
PT - useful in gene therapy for promoting and inhibiting apoptosis and
PT for diagnosing cells with potential for apoptosis
XX
PS Claim 1; Page 30-32; 48pp; English.
XX
CC A cDNA clone (T66970) includes to coding sequence for the proenzyme
CC (W15247) of CMH-1, a human cysteine protease involved in apoptosis.
CC To obtain the clone, the GenBank Expressed Sequence Tag (EST)
CC database was screened for clones encoding polypeptides similar to
CC interleukin-1b converting enzyme. Partial cDNA clone T50828 was
CC identified. Partial cDNA clone 72778 corresponding to this EST was
CC then obtd. A spleen cDNA pool was subjected to PCR using primers
CC (T66971-92) based on a 5' adaptor and EST T50828. A PCR fragment
CC contg. the 5' coding region was ligated to the clone 72778 3' region
CC to obtain the full-length CHM-1 sequence. CHM-1 (i.e. CPP32/Mch2
CC homologue-1) DNA can be used in gene therapy to promote apoptosis,
CC esp. in cancer cells, and to modulate patient cholesterol levels,
CC as well as to diagnose cells with potential for apoptosis.
CC Antisense sequences or sequences encoding inactive CHM-1 mutants
CC can be used to inhibit apoptosis, e.g. to prevent neurological
CC disease.
XX
SQ Sequence 2341 BP; 702 A; 455 C; 536 G; 648 T; 0 other;

Query Match 76.8%; Score 1052.8; DB 18; Length 2341;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1076; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

```
Qy   292 CGGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATT 351
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     7 cggagagagactgtgccagtcccagccgccctaccgccgtgggaacgatggcagatga-t 65

Qy   352 CAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCT 411
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    66 cagggctgtattgaggagcaggggttgaggattcagcaaatgaagattcagtggatgct 125

Qy   412 AAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACC 471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   126 aagccagaccggtcctcgtttgtaccgtccctcttcagtaagaagaagaaaaatgtcacc 185

Qy   472 ATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTT 531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   186 atgcgatccatcaagaccacccgggaccgagtgcctacatatcagtacaacatgaatttt 245

Qy   532 GAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATG 591
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   246 gaaaagctgggcaaatgcatcataataaacaacaagaactttgataaagtgacaggtatg 305

Qy   592 GGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTG 651
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   306 ggcgttcgaaacggaacagacaaagatgccgaggcgctcttcaagtgcttccgaagcctg 365

Qy   652 GGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAA 711
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   366 ggttttgacgtgattgtctataatgactgctcttgtgccaagatgcaagatctgcttaaa 425

Qy   712 AAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCAT 771
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   426 aaagcttctgaagaggaccatacaaatgccgcctgcttcgcctgcatcctcttaagccat 485

Qy   772 GGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCC 831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   486 ggagaagaaaatgtaatttatgggaaagatggtgtcacaccaataaaggatttgacagcc 545

Qy   832 CACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAG 891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   546 cactttaggggggatagatgcaaaacccttttagagaaacccaaactcttcttcattcag 605

Qy   892 GCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGAC 951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   606 gcttgccgagggaccgagcttgatgatggcatccaggccgactcggggcccatcaatgac 665

Qy   952 ACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCC 1011
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   666 acagatgctaatcctcgatacaagatcccagtggaagctgacttcctcttcgcctattcc 725

Qy  1012 ACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCC 1071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   726 acggttccaggctattactcgtggaggagcccaggaagaggctcctggtttgtgcaagcc 785

Qy  1072 CTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAG 1131
             ||||||||||||||||||||||||||||||||||||||||||||| ||||| ||||||
Db   786 ctctgctccatcctggaggagcacggaaaagacctggaaatcatgcagatcct-caccag 844

Qy  1132 GGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGA 1191
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   845 ggtgaatgacagagttgccaggcactttgagtctcagtctgatgacccacacttccatga 904

Qy  1192 GAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATA 1251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   905 gaagaagcagatcccctgtgtggtctccatgctcaccaaggaactctacttcagtcaata 964

Qy  1252 GCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATT 1311
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   965 gccatatcaggggtacattctagctgagaagcaatgggtcactcattaatgaatcacatt 1024

Qy  1312 TTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1025 tttttatgctcttgaaatattcagaaattctccaggattttaatttcaggaaaatgtatt 1084
```

RESULT 3
AAT66992
ID AAT66992 standard; cDNA; 2309 BP.
XX
AC AAT66992;
XX
DT 05-AUG-1997 (first entry)
XX
DE Apototic protease Mch3-alpha cDNA.
XX
KW Mch3-alpha; cysteine protease; apoptosis; AIDS; ischaemia;
KW neurodegenerative disease; gene therapy; diagnosis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 44..955
FT /*tag= a
FT intron 595..596
FT /*tag= b
FT /note= "site of intron in genomic DNA"
XX
PN WO9718313-A1.
XX
PD 22-MAY-1997.
XX
PF 12-NOV-1996; 96WO-US18118.
XX
PR 13-NOV-1995; 95US-0556627.
XX
PA (IDUN-) IDUN PHARM INC.

PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Alnemri ES, Armstrong R, Fernandes-Alnemri T, Litwack G;
PI Tomaselli K;
XX
DR WPI; 1997-289289/26.
DR P-PSDB; AAW15262.
XX
PT New gene encoding Mch3, a cysteine protease that regulates apoptosis
PT - for treating human diseases associated with apoptosis, and
PT screening for antagonists and agonists of Mch3
XX
PS Claim 2; Fig 1; 52pp; English.
XX
CC A cDNA clone (AAT66992) codes for Mch3-alpha (AAW15262), an apoptotic
CC protein and novel member of the interleukin-1-beta converting
CC enzyme (ICE) family of cysteine proteases. It was obtd. by
CC screening a human database of expressed sequence tags for sequences
CC showing homology to ICE family members, and use of a partial cDNA
CC clone with high homology to CPP32 and Ced-3 to screen a Jurkat cDNA
CC library. A second cDNA clone, Mch3-beta (AAT66993), contains a
CC deletion and insertion corresponding to nucleotides 488-592 of the
CC Mch3-alpha sequence. Mch3 nucleic acids can be used to produce
CC recombinant Mch3 polypeptides useful for treating diseases
CC associated with cell death, such as AIDS, ischaemic injury,
CC neurodegenerative diseases, etc. They can also be used in gene
CC therapy to regulate apoptosis.
XX
SQ Sequence 2309 BP; 680 A; 453 C; 531 G; 645 T; 0 other;

Query Match 76.6%; Score 1050.4; DB 18; Length 2309;
Best Local Similarity 99.7%; Pred. No. 8.2e-315;
Matches 1073; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```
Qy    296 GAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGG 355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db      1 gagagactgtgccagtcccagccgccctaccgccgtgggaacgatggcagatga-tcagg 59

Qy    356 GCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGC 415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 gctgtattgaagagcaggggttgaggattcagcaaatgaagattcagtggatgctaagc 119

Qy    416 CAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGC 475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 cagaccggtcctcgtttgtaccgtccctcttcagtaagaagaagaaaaatgtcaccatgc 179

Qy    476 GATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAA 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 gatccatcaagaccacccgggaccgagtgcctacatatcagtacaacatgaattttgaaa 239

Qy    536 AGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCG 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 agctgggcaaatgcatcataataaacaacaagaactttgataaagtgacaggtatgggcg 299

Qy    596 TTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTT 655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 ttcgaaacggaacagacaaagatgccgaggcgctcttcaagtgcttccgaagcctgggtt 359

Qy    656 TTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAG 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 ttgacgtgattgtctataatgactgctcttgtgccaagatgcaagatctgcttaaaaaag 419

Qy    716 CTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAG 775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 cttctgaagaggaccatacaaatgccgcctgcttcgcctgcatcctcttaagccatggag 479

Qy    776 AAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACT 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 aagaaaatgtaatttatgggaaagatggtgtcacaccaataaaggatttgacagcccact 539

Qy    836 TTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTT 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 ttaggggggatagatgcaaaaccctttttagagaaacccaaactcttcttcattcaggctt 599

Qy    896 GCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAG 955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 gccgagggaccgagcttgatgatggcatccaggccgactcggggcccatcaatgacacag 659

Qy    956 ATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGG 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 atgctaatcctcgatacaagatcccagtggaagctgacttcctcttcgcctattccacgg 719

Qy   1016 TTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCT 1075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 ttccaggctattactcgtggaggagcccaggaagaggctcctggtttgtgcaagccctct 779

Qy   1076 GCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTG 1135
          |||||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||
Db    780 gctccatcctggaggagcacggaaaagacctggaaatcatgcagatcct-caccagggtg 838

Qy   1136 AATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAG 1195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    839 aatgacagagttgccaggcactttgagtctcagtctgatgacccacacttccatgagaag 898

Qy   1196 AAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCA 1255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    899 aagcagatcccctgtgtggtctccatgctcaccaaggaactctacttcagtcaatagcca 958

Qy   1256 TATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTT 1315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    959 tatcaggggtacattctagctgagaagcaatgggtcactcattaatgaatcacatttttt 1018

Qy   1316 TATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1019 tatgctcttgaaatattcagaaattctccaggattttaatttcaggaaaatgtatt 1074
```

RESULT 4
AAD03914
ID AAD03914 standard; DNA; 2309 BP.
XX
AC AAD03914;
XX
DT 02-JUL-2001 (first entry)
XX
DE Human caspase-7 DNA.
XX
KW Human; caspase-7; interleukin-1 converting enzyme; ICE-LAP3; CHM-1; Mhc3;
KW cysteine protease; apoptosis; caspase expression cassette; metastasis;
KW tumour; cathepsin B; urokinase; proliferation; gene therapy; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 44..955
FT /*tag= a
FT /product= "Human caspase-7"
XX
PN WO200129232-A2.
XX
PD 26-APR-2001.
XX
PF 19-OCT-2000; 2000WO-US28941.
XX
PR 20-OCT-1999; 99US-0160559.
PR 14-AUG-2000; 2000US-0225564.
XX
PA (SCIO-) SCIOS INC.
XX
PI Cordell B, Li Y;
XX
DR WPI; 2001-290920/30.
DR P-PSDB; AAE00604.
XX
PT Novel fusion polypeptide comprising first and second caspase subunit

```
PT   separated by cleavage site not associated in nature with caspase
PT   subunit, useful for cloning gene encoding enzymes involved in
PT   proteolytic cleavage  -
XX
PS   Disclosure; Fig 13; 116pp; English.
XX
CC   The present sequence is a DNA encoding human Caspase-7 also known as
CC   interleukin-1 converting enzyme (ICE) LAP3, CHM-1 and Mhc3. Caspases are
CC   a family of cysteine proteases, that participate in the initiation and
CC   execution of apoptosis. Caspases exist as pro-enzymes, activated by
CC   cleavage into a large and small subunit, occurring after specific
CC   aspartic acid residues within the pro-enzyme sequence.
CC   The present invention relates to a method for functional cloning of genes
CC   encoding proteins or enzymes involved in proteolytic cleavage. The
CC   invention is based on the use of caspase expression cassettes comprising
CC   the coding sequence of a proteolytic cleavage site flanked by sequences
CC   encoding two caspase subunits. A fusion polypeptide comprising a first
CC   and a second caspase subunit, separated by a cleavage site not associated
CC   in nature, is useful for cloning gene encoding enzymes involved in
CC   proteolytic cleavage. An expression cassette containing fusion
CC   polypeptide is used to identify a mutant cell line deficient in an
CC   enzyme of interest and is also useful for diagnosis and suppression of
CC   proliferation or metastases of a tumour cell characterised by
CC   overexpression of a polypeptide (e.g. Cathepsin B or urokinase,
CC   selectively expressed in the tumour cells). DNA encoding fusion
CC   polypeptide is used in gene therapy.
XX
SQ   Sequence 2309 BP; 680 A; 453 C; 531 G; 645 T; 0 other;
```

```
  Query Match             76.6%;  Score 1050.4;  DB 22;  Length 2309;
  Best Local Similarity   99.7%;  Pred. No. 8.2e-315;
  Matches 1073;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy     296 GAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGG 355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db       1 gagagactgtgccagtcccagccgccctaccgccgtgggaacgatggcagatga-tcagg 59

Qy     356 GCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGC 415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60 gctgtattgaagagcaggggttgaggattcagcaaatgaagattcagtggatgctaagc 119

Qy     416 CAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGC 475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     120 cagaccggtcctcgtttgtaccgtccctcttcagtaagaagaagaaaaatgtcaccatgc 179

Qy     476 GATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAA 535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180 gatccatcaagaccacccgggaccgagtgcctacatatcagtacaacatgaattttgaaa 239

Qy     536 AGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCG 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     240 agctgggcaaatgcatcataataaacaacaagaactttgataaagtgacaggtatgggcg 299

Qy     596 TTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTT 655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 ttcgaaacggaacagacaaagatgccgaggcgctcttcaagtgcttccgaagcctgggtt 359

Qy     656 TTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAG 715
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 ttgacgtgattgtctataatgactgctcttgtgccaagatgcaagatctgcttaaaaaag 419

Qy     716 CTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAG 775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     420 cttctgaagaggaccatacaaatgccgcctgcttcgcctgcatcctcttaagccatggag 479

Qy     776 AAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACT 835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     480 aagaaaatgtaatttatgggaaagatggtgtcacaccaataaaggatttgacagcccact 539

Qy     836 TTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTT 895
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     540 ttaggggggatagatgcaaaacccttttagagaaacccaaactcttcttcattcaggctt 599

Qy     896 GCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAG 955
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 gccgagggaccgagcttgatgatggcatccaggccgactcggggcccatcaatgacacag 659

Qy     956 ATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGG 1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     660 atgctaatcctcgatacaagatcccagtggaagctgacttcctcttcgcctattccacgg 719

Qy    1016 TTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCT 1075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     720 ttccaggctattactcgtggaggagcccaggaagaggctcctggtttgtgcaagccctct 779

Qy    1076 GCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTG 1135
           |||||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||
Db     780 gctccatcctggaggagcacggaaaagacctggaaatcatgcagatcct-caccagggtg 838

Qy    1136 AATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAG 1195
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     839 aatgacagagttgccaggcactttgagtctcagtctgatgacccacacttccatgagaag 898

Qy    1196 AAGCAGATCCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCA 1255
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     899 aagcagatcccctgtgtggtctccatgctcaccaaggaactctacttcagtcaatagcca 958

Qy    1256 TATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTT 1315
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     959 tatcaggggtacattctagctgagaagcaatgggtcactcattaatgaatcacatttttt 1018

Qy    1316 TATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1019 tatgctcttgaaatattcagaaattctccaggattttaatttcaggaaaatgtatt 1074
```

```
RESULT   5
AAX81223
ID   AAX81223 standard; DNA; 913 BP.
XX
AC   AAX81223;
XX
DT   10-SEP-1999  (first entry)
XX
DE   Nucleotide sequence of caspase-7 (Mch3).
XX
KW   Rev-caspase; cysteine protease; zymogen; caspase; cancer; gene therapy;
KW   autoimmune disease; caspase-mediated apoptosis; neurodegenerative;
KW   tumour cell; myocardial infarction; human; ds.
XX
OS   Homo sapiens.
XX
PN   WO9935277-A2.
XX
PD   15-JUL-1999.
XX
PF   11-JAN-1999;   99WO-US00632.
XX
PR   09-JAN-1998;   98US-0070987.
XX
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES;
XX
DR   WPI; 1999-419353/35.
DR   P-PSDB; AAY21721.
XX
PT   New isolated nucleic acid molecule encoding a rev-caspase  - used
PT   for screening and identifying inhibitors or enhancers for treating
PT   cancer or autoimmune disease
XX
PS   Disclosure; Fig 17A-B; 74pp; English.
XX
CC   The invention relates to nucleic acid molecules encoding rev-caspases.
CC   Rev-caspases are cysteine proteases that specifically cleave proteins
```

```
CC   after Asp residues and is expressed as a zymogen, in which the small
CC   subunit is N-terminal to a large subunit. A gene delivery vehicle
CC   comprising a rev-caspase coding sequence is useful for the treatment of
CC   cancer, where the gene delivery vehicle is internalised by tumour cells.
CC   The gene delivery vehicle can also be used to treat autoimmune diseases.
CC   Cells transfected with a rev-caspase expressing vector can be used in
CC   identification of inhibitors or enhancers of caspase-mediated apoptosis.
CC   In vitro translated rev-caspase can be used to identify an inhibitor or
CC   enhancer of caspase processing activity. Caspase inhibitors are useful
CC   for treating neurodegenerative diseases as well as for inhibiting
CC   apoptosis in the heart following myocardial infarction. Sequences
CC   AAX81217 -AAX81226 represent human caspase genes encoding caspase 1-10
CC   gene products (AAY21715-Y21724).
XX
SQ   Sequence 913 BP; 263 A; 218 C; 224 G; 207 T; 1 other;

  Query Match             64.8%;  Score 888.4;  DB 20;  Length 913;
  Best Local Similarity   99.8%;  Pred. No. 6.3e-265;
  Matches  900;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy    351 TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGC 410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     12 tcagggctgtattgaagagcaggggttgaggattcagcaaatgaagattcagtggatgc 71

Qy    411 TAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCAC 470
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     72 taagccagaccggtcctcgtttgtaccgtccctcttcagtaagaagaagaaaaatgtcac 131

Qy    471 CATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTT 530
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    132 catgcgatccatcaagaccacccgggaccgagtgcctacatatcagtacaacatgaattt 191

Qy    531 TGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTAT 590
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    192 tgaaaagctgggcaaatgcatcataataaacaacaagaactttgataaagtgacaggtat 251

Qy    591 GGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCT 650
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    252 gggcgttcgaaacggaacagacaaagatgccgaggcgctcttcaagtgcttccgaagcct 311

Qy    651 GGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAA 710
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    312 gggttttgacgtgattgtctataatgactgctcttgtgccaagatgcaagatctgcttaa 371

Qy    711 AAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCA 770
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    372 aaaagcttctgaagaggaccatacaaatgccgcctgcttcgcctgcatcctcttaagcca 431

Qy    771 TGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGC 830
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    432 tggagaagaaaatgtaatttatgggaaagatggtgtcacaccaataaaggatttgacagc 491

Qy    831 CCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCA 890
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 ccactttaggggggatagatgcaaaacccttttagagaaacccaaactcttcttcattca 551

Qy    891 GGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGA 950
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    552 ggcttgccgagggaccgagcttgatgatggcatccaggccgactcggggcccatcaatga 611

Qy    951 CACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTC 1010
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    612 cacagatgctaatcctcgatacaagatcccagtggaagctgacttcctcttcgcctattc 671

Qy   1011 CACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGC 1070
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    672 cacggttccaggctattactcgtggaggagcccaggaagaggctcctggtttgtgcaagc 731

Qy   1071 CCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCA 1130
          |||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db    732 cctctgctccatcctggaggagcacggaaaagacctggaaatcatgcagatcct-cacca 790
```

```
Qy   1131 GGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATG 1190
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    791 gggtgaatgacagagttgccaggcactttgagtctcagtctgatgacccacacttccatg 850

Qy   1191 AGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAAT 1250
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    851 agaagaagcagatcccctgtgtggtctccatgctcaccaaggaactctacttcagtcaat 910

Qy   1251 AG 1252
          ||
Db    911 ag 912


RESULT   6
AAT66993
ID   AAT66993 standard; cDNA; 2377 BP.
XX
AC   AAT66993;
XX
DT   05-AUG-1997  (first entry)
XX
DE   Cysteine protease Mch3-beta cDNA.
XX
KW   Mch3-beta; cysteine protease; apoptosis; AIDS; ischaemia; .
KW   neurodegenerative disease; gene therapy; diagnosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             146..839
FT                   /*tag= a
XX
PN   WO9718313-A1.
XX
PD   22-MAY-1997.
XX
PF   12-NOV-1996;   96WO-US18118.
XX
PR   13-NOV-1995;   95US-0556627.
XX
PA   (IDUN-) IDUN PHARM INC.
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES,  Armstrong R,  Fernandes-Alnemri T,  Litwack G;
PI   Tomaselli K;
XX
DR   WPI; 1997-289289/26.
DR   P-PSDB; AAW15263.
XX
PT   New gene encoding Mch3, a cysteine protease that regulates apoptosis
PT   - for treating human diseases associated with apoptosis, and
PT   screening for antagonists and agonists of Mch3
XX
PS   Claim 2; Fig 1; 52pp; English.
XX
CC   A cDNA clone (AAT66993) codes for Mch3-beta (AAW15263), a novel member
CC   of the interleukin-1-beta converting enzyme (ICE) family of
CC   cysteine proteases.  It was obtd. by screening a human database of
CC   expressed sequence tags for sequences showing homology to ICE
CC   family members, and use of a partial cDNA clone with high homology
CC   to CPP32 and Ced-3 to screen a Jurkat cDNA library.  Mch3-beta cDNA
CC   contains a deletion and insertion corresponding to nucleotides
CC   488-592 of Mch3-alpha cDNA (AAT66992) and also has a longer 3'
CC   nontranslated sequence.  Mch3-beta lacks the active site of Mch3-
CC   alpha.  Mch3 nucleic acids can be used to produce recombinant Mch3
CC   polypeptides useful for treating diseases associated with cell
CC   death, such as AIDS, ischaemic injury, neurodegenerative diseases,
CC   etc.  They can also be used in gene therapy to regulate apoptosis.
XX
SQ   Sequence 2377 BP; 680 A; 485 C; 565 G; 647 T; 0 other;

  Query Match             62.6%;  Score 858;  DB 18;  Length 2377;
```

```
  Best Local Similarity   92.6%;  Pred. No. 3.2e-255;
  Matches  958;  Conservative    0;  Mismatches   40;  Indels   36;  Gaps    4;

Qy     338 GATGGCAGATGATTCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAG 397
           ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db     145 gatggcagatga-tcagggctgtattgaagagcagggggttgaggattcagcaaatgaag 203

Qy     398 ATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGA 457
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     204 attcagtggatgctaagccagaccggtcctcgtttgtaccgtccctcttcagtaagaaga 263

Qy     458 AGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGT 517
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     264 agaaaaatgtcaccatgcgatccatcaagaccacccgggaccgagtgcctacatatcagt 323

Qy     518 ACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATA 577
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     324 acaacatgaattttgaaaagctgggcaaatgcatcataataaacaacaagaactttgata 383

Qy     578 AAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGT 637
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     384 aagtgacaggtatgggcgttcgaaacggaacagacaaagatgccgaggcgctcttcaagt 443

Qy     638 GCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGC 697
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     444 gcttccgaagcctgggttttgacgtgattgtctataatgactgctcttgtgccaagatgc 503

Qy     698 AAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCA 757
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     504 aagatctgcttaaaaaagcttctgaagaggaccatacaaatgccgcctgcttcgcctgca 563

Qy     758 TCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAA 817
           |||||||||||||||||||||||| |        ||||   || | |  |||| |
Db     564 tcctcttaagccatggagaagaaaatat--------ggaatcttgctctgtcacccaggc 615

Qy     818 AGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAAC 877
           ||| |  |  | |    |   |                          ||| |   |
Db     616 tggagtgcagcggcgtgatctcg------------------------gaagactgcaa 649

Qy     878 TCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGG 937
           || |  |   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     650 cctccacctcccaggcttgccgagggaccgagcttgatgatggcatccaggccgactcgg 709

Qy     938 GGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCC 997
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     710 ggcccatcaatgacacagatgctaatcctcgatacaagatcccagtggaagctgacttcc 769

Qy     998 TCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCT 1057
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     770 tcttcgcctattccacggttccaggctattactcgtggaggagcccaggaagaggctcct 829

Qy    1058 GGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGC 1117
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     830 ggtttgtgcaagccctctgctccatcctggaggagcacggaaaagacctggaaatcatgc 889

Qy    1118 AAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGAC 1177
           | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     890 agatcct-caccagggtgaatgacagagttgccaggcactttgagtctcagtctgatgac 948

Qy    1178 CCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTC 1237
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     949 ccacacttccatgagaagaagcagatcccctgtgtggtctccatgctcaccaaggaactc 1008

Qy    1238 TACTTCAGTCAATAGCCATATCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCAT 1297
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1009 tacttcagtcaatagccatatcaggggtacattctagctgagaagcaatgggtcactcat 1068

Qy    1298 TAATGAATCACATTTTTTTATGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTT 1357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1069 taatgaatcacatttttttatgctcttgaaatattcagaaattctccaggattttaattt 1128

Qy    1358 CAGGAAAATGTATT 1371
           ||||||||||||||
Db    1129 caggaaaatgtatt 1142


RESULT   7
AAS10510
ID   AAS10510 standard; DNA; 1297 BP.
XX
AC   AAS10510;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Mouse caspase 3 DNA sequence.
XX
KW   Mouse; caspase 3; apoptosis; hyperproliferative disorder; hepatitis;
KW   viral infection; haematopoietic disorder; autoimmune disorder;
KW   atherosclerosis; neurological disorder; ds.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             68..901
FT                   /*tag= a
FT                   /product= "Caspase 3"
XX
PN   WO200153310-A1.
XX
PD   26-JUL-2001.
XX
PF   11-JAN-2001; 2001WO-US00888.
XX
PR   18-JAN-2000; 2000US-0484617.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Zhang H,  Cowsert LM;
XX
DR   WPI; 2001-442252/47.
DR   P-PSDB; AAU05395.
XX
PT   New antisense compound to inhibit caspase 3 is useful for treating
PT   hepatitis and atherosclerosis -
XX
PS   Example 13; Page 96-97; 127pp; English.
XX
CC   The present sequence encoding for mouse caspase 3 is described
CC   in an invention relating to novel antisense oligonucleotides
CC   (AAS10517-AAS10676) and methods of using these compounds for the
CC   modulation of caspase 3 expression. The caspase 3 antisense
CC   oligonucleotides specifically hybridise with and inhibit the
CC   expression of caspase 3. Antisense compounds targeted to caspase 3
CC   are useful to inhibit caspase 3 expression in cells or tissues and
CC   to modulate apoptosis. The caspase 3 antisense oligonucleotides are
CC   useful for treating disorders associated with expression of caspase 3.
CC   Such disorders include hyperproliferative disorders (e.g. cancer),
CC   viral infections (e.g. hepatitis), haematopoietic disorders, autoimmune
CC   disorders, atherosclerosis and neurological disorders (e.g. Alzheimer's
CC   disease).
XX
SQ   Sequence 1297 BP; 380 A; 255 C; 314 G; 348 T; 0 other;


  Query Match           17.0%;  Score 232.8;  DB 22;  Length 1297;
  Best Local Similarity   59.0%;  Pred. No. 2.5e-61;
  Matches  441;  Conservative    0;  Mismatches  297;  Indels   10;  Gaps    2;

Qy     517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
           ||||| ||| ||| |     |  |||||| |||||| ||||| || || ||||||||  ||
Db     176 tacaaaatggattatcctgaaatgggcatatgcataataattaataataagaacttccat 235

Qy     577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
           ||    || || |||     |||    || || ||    || || |     |||    ||
Db     236 aagagcactggaatgtcatctcgctctggtacggatgtggacgcagccaacctcagagag 295
```

```
Qy    637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
Db    296 acattcatgggcctgaaataccaagtcaggaataaaaatgatcttactcgtgaagacatt 355

Qy    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
Db    356 ttggaattaatggatagtgtttctaaggaagatcatagcaaaaggagcagctttgtgtgt 415

Qy    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
Db    416 gtgattctaagccatggtgatgaaggggtcatttatgggacaaatgggcctgttgaactg 475

Qy    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
Db    476 aaaaagttgactagcttcttcagaggcgactactgccggagtctgactggaaagccgaaa 535

Qy    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
Db    536 ctcttcatcattcaggcctgccggggtacggagctggactgtggcattgagacagacagt 595

Qy    937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
Db    596 gggactgatgaggagatgg---------cttgccagaagataccggtggaggctgacttc 646

Qy    997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
Db    647 ctgtatgcttactctacagcacctggttactattcctggagaaattcaaaggacgggtcg 706

Qy   1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
Db    707 tggttcatccagtccctttgcagcatgctgaagctgtacgcgcacaagctagaatttatg 766

Qy   1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
Db    767 cacattct-cactcgcgttaacaggaaggtggcaacggaattcgagtccttctccctgga 825

Qy   1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
Db    826 ctccactttccacgcaaagaaacagatcccgtgtattgtgtccatgctcacgaaagaact 885

Qy   1237 CTACTTCAGTCAATAGCCATATCAGGGG 1264
Db    886 gtacttttatcactagaggaatgattgg 913
```

RESULT 8
AAC64043
ID AAC64043 standard; DNA; 2538 BP.
XX
AC AAC64043;
XX
DT 14-FEB-2001 (first entry)
XX
DE pYNGCas3p10p35 insert encoding human caspase-3 and AcNPV p35 protein.
XX
KW Human caspase-3; AcNPV p35 protein; BmNPV p10 promoter;
KW polyhedron promoter; baculovirus vector; apoptosis inhibitor protein;
KW recominant caspase production; insect cell; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Autographa californica nuclear polyhedrosis virus.
OS Chimeric - Bombyx mori nuclear polyhedrosis virus.
XX
PN JP2000253882-A.
XX
PD 19-SEP-2000.
XX
PF 10-MAR-1999; 99JP-0062732.
XX
PR 10-MAR-1999; 99JP-0062732.
XX
PA (KATA) KATAKURA IND CO LTD.
XX
DR WPI; 2000-658811/64.
XX
PT Baculovirus vector used for the preparation of caspase protein,
PT comprises a transfer vector plasmid containing a caspase
PT protein-producing gene -
XX
PS Example 1; Page 7-8; 10pp; Japanese.
XX
CC The invention relates to a novel baculovirus vector used for the
CC preparation of a caspase protein, and containing a recombinant transfer
CC vector plasmid comprising a gene encoding a caspase and a gene encoding
CC an apoptosis-inhibiting protein. The invention also encompasses the
CC preparation of caspase protein, comprising transfecting the novel
CC baculovirus vector into an insect or an insect cell; breeding the
CC insect or culturing the cell; and collecting the caspase protein from
CC the insect or insect cell. The present sequence represents the expression
CC cassette insert from a transfer vector (pYNGCas3p10p35) produced in an
CC exemplification of the invention. This insert comprises a mature human
CC caspase-3 gene under the control of a polyhedron promoter, and an
CC Autographa californica nuclear polyhedrosis virus (AcNPV) p35 gene under
CC the control of a Bombyx mori nuclear polyhedrosis virus (BmNPV) p10
CC promoter.
XX
SQ Sequence 2538 BP; 865 A; 446 C; 506 G; 721 T; 0 other;

Query Match 16.8%; Score 230.6; DB 21; Length 2538;
Best Local Similarity 58.7%; Pred. No. 1.8e-60;
Matches 440; Conservative 0; Mismatches 299; Indels 10; Gaps 2;

```
Qy    517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
Db    244 tataaaatggattatcctgagatgggtttatgtataataattaataataagaattttcat 303

Qy    577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
Db    304 aaaagcactggaatgacatctcggtctggtacagatgtcgatgcagcaaacctcagggaa 363

Qy    637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
Db    364 acattcagaaacttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaaatt 423

Qy    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
Db    424 gtggaattgatgcgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtttgt 483

Qy    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
Db    484 gtgcttctgagccatggtgaagaaggaataatttttggaacaaatggacctgttgacctg 543

Qy    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
Db    544 aaaaaaataacaaacttttcagaggggatcgttgtagaagtctaactggaaaacccaaa 603

Qy    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
Db    604 cttttcattattcaggcctgccgtggtacagaactggactgtggcattgagacagacagt 663

Qy    937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
Db    664 gg---------tgttgatgatgacatggcgtgtcataaaataccagtggaggccgacttc 714

Qy    997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
Db    715 ttgtatgcatactccacagcacctggttattattcttggcgaaattcaaaggatggctcc 774

Qy   1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
Db    775 tggttcatccagtcgctttgtgccatgctgaaacagtatgccgacaagcttgaatttatg 834

Qy   1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
Db    835 cacattct-tacccgggttaaccgaaaggtggcaacagaatttgagtccttttccttga 893
```

```
Qy   1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
Db    894 cgctacttttcatgcaaagaaacagattccatgtattgtttccatgctcacaaaagaact 953

Qy   1237 CTACTTCAGTCAATAGCCATATCAGGGGT 1265
Db    954 ctatttttatcactaagaattccatggat 982
```

RESULT 9
AAI66510
ID AAI66510 standard; DNA; 834 BP.
XX
AC AAI66510;
XX
DT 11-DEC-2001 (first entry)
XX
DE Pig caspase coding sequence #2.
XX
KW Pig; caspase; inflammation; infection; ds.
XX
OS Sus scrofa.
XX
PN JP2001169785-A.
XX
PD 26-JUN-2001.
XX
PF 03-OCT-2000; 2000JP-0303931.
XX
PR 05-OCT-1999; 99JP-0284829.
XX
PA (NORQ) NORINSUISANSHO KACHIKU EISEI.
PA (MUNE/) MUNETA Y.
PA (MORI/) MORI Y.
PA (SHIM/) SHIMOCHI Y.
PA (ARAI/) ARAI K.
XX
DR WPI; 2001-592539/67.
DR P-PSDB; AAG78712.
XX
PT Structural gene of swine caspase family, used for artificial control of
PT inflammatory reactions in infectious diseases -
XX
PS Example 2; Page 15-16; 23pp; Japanese.
XX
CC The present invention provides two versions of a porcine caspase protein.
CC These can be used to control inflammatory reactions to microbial or viral
CC infectious diseases. The present sequence is the coding sequence of one
CC version of the protein of the invention.
XX
SQ Sequence 834 BP; 276 A; 149 C; 179 G; 230 T; 0 other;

Query Match 16.8%; Score 230.4; DB 22; Length 834;
Best Local Similarity 59.1%; Pred. No. 1e-60;
Matches 435; Conservative 0; Mismatches 291; Indels 10; Gaps 2;

```
Qy    517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
Db    109 tacaaaatggattatcctgaaatgggtttatgtataataattaataataagaactttgat 168

Qy    577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
Db    169 aaaaataccggaatggcatgtcgatctggtacagatgtggatgctgcaaatctcagggag 228

Qy    637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
Db    229 accttcacaaacttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaaatt 288

Qy    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
Db    289 ttggagttaatgcacagtgtttctaaagaagaccatagcaaaaggagcagttttatttgc 348

Qy    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
Db    349 gtgcttctaagccatggtgaagaaggaaaaatttttggaacaaatggacctgttgatctg 408

Qy    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
Db    409 aaaaaattaacaagtttcttcagaggggactgctgtagaactctaactggcaaacccaaa 468

Qy    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
Db    469 cttttcataattcaggcctgccgaggcacagaattggactgtgggattgagacggacagt 528

Qy    937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
Db    529 gggactgaagatgacatgg---------cgtgtcagaaaataccagttgaggcagacttc 579

Qy    997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
Db    580 ttgtatgcatattctacagcacctggttactattcctggcgaaattcaaaggacggatcc 639

Qy   1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
Db    640 tggttcatccagtcactttgtgcagcgctgaaacagtatgttcacaagcttgagcttatg 699

Qy   1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
Db    700 cacattct-tactcgggttaaccgaaaggtagcagtagaattcgagtccttttctactga 758

Qy   1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
Db    759 ctctacttttcatgcaaagaaacagattccatgtattgtgtccatgctcacaaaagaact 818

Qy   1237 CTACTTCAGTCAATAG 1252
Db    819 ctacttttatcactag 834
```

RESULT 10
AAT33567
ID AAT33567 standard; cDNA; 1158 BP.
XX
AC AAT33567;
XX
DT 06-DEC-1996 (first entry)
XX
DE Pro-Yama cDNA.
XX
KW Yama; CrmA; apoptosis; Fas receptor; T-lymphocyte; AIDS; HIV;
KW human immunodeficiency virus; gene therapy; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 224..1057
FT /*tag= a
XX
PN WO9625501-A2.
XX
PD 22-AUG-1996.
XX
PF 09-FEB-1996; 96WO-US01882.
XX
PR 01-JUN-1995; 95US-0457731.
PR 13-FEB-1995; 95US-0389812.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Dixit VM;
XX
DR WPI; 1996-393402/39.
DR P-PSDB; AAW00677.
XX

PT New nucleic acid encoding Yama protein or CrmA mutant - useful for
PT modulating apoptosis, maintaining T cell viability in AIDS patients
PT and for drug screening
XX
PS Disclosure; Fig 1; 118pp; English.
XX
CC A cDNA clone (AAT33567) codes for pro-Yama (AAW00677), a zymogen which
CC upon activation cleaves PARP to an 85 kDa form. Activated Yama has
CC the ability to modulate cellular function associated with the Fas
CC receptor pathway such as Fas-associated apoptosis. This activity is
CC inhibited by CrmA (see also AAW00678). The cDNA was isolated from a
CC cDNA library prepd. from human umbilical vein endothelial cells
CC treated with tumour necrosis factor and cycloheximide. It can be
CC used to produce Yama in a host cell for use in drug screening.
CC Yama and anti-Yama protein nucleic acids, or antisense RNA, can be
CC used to modulate cellular functions such as apoptosis and immune
CC disorders mediated by the Fas pathway.
XX
SQ Sequence 1158 BP; 328 A; 220 C; 275 G; 335 T; 0 other;

Query Match 16.7%; Score 229; DB 17; Length 1158;
Best Local Similarity 58.6%; Pred. No. 3.5e-60;
Matches 439; Conservative 0; Mismatches 300; Indels 10; Gaps 2;

Qy 517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
Db 332 tataaaatggattatcctgagatgggtttatgtataataattaataataagaattttcat 391

Qy 577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
Db 392 aaaagcactggaatgacatctcggtctggtacagatgtcgatgcagcaaacctcagggaa 451

Qy 637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
Db 452 acattcagaaacttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaaatt 511

Qy 697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
Db 512 gtggaattgatgcgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtttgt 571

Qy 757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
Db 572 gtgcttctgagccatggtgaagaaggaataatttttggaacaaatggacctgttgacctg 631

Qy 817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
Db 632 aaaaaaataacaaacttttcagaggggatcgttgtagaagtctaactggaaaacccaaa 691

Qy 877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
Db 692 cttttcattattcaggcctgccgtggtacagaactggactgtggcattgagacagacagt 751

Qy 937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
Db 752 gg---------tgttgatgatgacatggcgtgtcataaaataccagtggaggccgacttc 802

Qy 997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
Db 803 ttgtatgcatactccacagcacctggttattattcttggcgaaattcaaaggatggctcc 862

Qy 1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
Db 863 tggttcatccagtcgctttgtgccatgctgaaacagtatgccgacaagcttgaatttatg 922

Qy 1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
Db 923 cacattct-tacccgggttaaccgaaaggtggcaacagaatttgagtccttttccttga 981

Qy 1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
Db 982 cgctacttttcatgcaaagaaacagattccatgtattgtttccatgctcacaaaagaact 1041

Qy 1237 CTACTTCAGTCAATAGCCATATCAGGGGT 1265
Db 1042 ctatttttatcactaaagaaatggttggt 1070

RESULT 11
AAV05471
ID AAV05471 standard; DNA; 1158 BP.
XX
AC AAV05471;
XX
DT 02-JUL-1998 (first entry)
XX
DE Nucleic acid encoding a protein designated YAMA.
XX
KW CrmA; tumour necrosis factor receptor; TNF-R; U1-70; inhibition;
KW apoptosis; activation; U1-70 path; apoptosis path; identification;
KW modulation; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 224..1057
FT /*tag= a
XX
PN JP09299077-A.
XX
PD 25-NOV-1997.
XX
PF 27-JAN-1997; 97JP-0012932.
XX
PR 09-FEB-1996; 96US-0591605.
XX
PA (UNMI) UNIV MICHIGAN.
XX
DR WPI; 1998-056551/06.
DR P-PSDB; AAW41688.
XX
PT Controlling apoptosis by inhibiting activation of the U1-70 pathway
PT - by introduction of nucleic acid encoding protein with CrmA
PT biological activity into the cell
XX
PS Disclosure; Fig 1; 48pp; Japanese.
XX
CC The present sequence encodes a non-naturally occurring protein designated
CC YAMA. The specification describes a novel method of controlling cell
CC functions controlled by the tumour necrosis factor receptor (TNF-R)
CC pathway containing U1-70. The method comprises introducing a nucleic
CC acid molecule encoding CrmA into a cell, and culturing the cell under
CC suitable conditions for transcription and translation of the CrmA nucleic
CC acid. Method for preventing and inhibiting apoptosis by inhibiting
CC activation of the U1-70 path, identifying the chemical related to the
CC apoptosis path, and screening a candidate chemical having biological
CC function in the apoptosis path are also described. The methods can be
CC used to identify substances which modulate apoptosis in cells controlled
CC by the TNF-R pathway.
XX
SQ Sequence 1158 BP; 328 A; 220 C; 275 G; 335 T; 0 other;

Query Match 16.7%; Score 229; DB 19; Length 1158;
Best Local Similarity 58.6%; Pred. No. 3.5e-60;
Matches 439; Conservative 0; Mismatches 300; Indels 10; Gaps 2;

Qy 517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
Db 332 tataaaatggattatcctgagatgggtttatgtataataattaataataagaattttcat 391

Qy 577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
Db 392 aaaagcactggaatgacatctcggtctggtacagatgtcgatgcagcaaacctcagggaa 451

Qy 637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696

```
Db    452 acattcagaaacttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaaatt 511

Qy    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756

Db    512 gtggaattgatgcgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtttgt 571

Qy    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816

Db    572 gtgcttctgagccatggtgaagaaggaataatttttggaacaaatggacctgttgacctg 631

Qy    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876

Db    632 aaaaaaataacaaacttttcagaggggatcgttgtagaagtctaactggaaaacccaaa 691

Qy    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936

Db    692 cttttcattattcaggcctgccgtggtacagaactggactgtggcattgagacagacagt 751

Qy    937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996

Db    752 gg---------tgttgatgatgacatggcgtgtcataaaataccagtggaggccgacttc 802

Qy    997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056

Db    803 ttgtatgcatactccacagcacctggttattattcttggcgaaattcaaaggatggctcc 862

Qy   1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116

Db    863 tggttcatccagtcgctttgtgccatgctgaaacagtatgccgacaagcttgaatttatg 922

Qy   1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176

Db    923 cacattct-tacccgggttaaccgaaaggtggcaacagaatttgagtccttttccttga 981

Qy   1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236

Db    982 cgctacttttcatgcaaagaaacagattccatgtattgtttccatgctcacaaaagaact 1041

Qy   1237 CTACTTCAGTCAATAGCCATATCAGGGGT 1265

Db   1042 ctatttttatcactaaagaaatggttggt 1070
```

```
RESULT 12
AAT15277
ID   AAT15277 standard; cDNA; 1159 BP.
XX
AC   AAT15277;
XX
DT   28-OCT-1996  (first entry)
XX
DE   Human interleukin-I-converting enzyme-like apoptosis protease-4 cDNA.
XX
KW   ICE-LAP-4; interleukin-I-converting enzyme-like apoptosis protease;
KW   enzyme; Alzheimer's disease; Parkinson's disease; septic shock;
KW   head injury; rheumatoid arthritis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             115..1137
FT                   /*tag= a
XX
PN   WO9613603-A1.
XX
PD   09-MAY-1996.
XX
PF   06-JUN-1995;   95WO-US07235.
XX
PR   01-NOV-1994;   94US-0334251.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Hastings GA, He W, Hudson PL, Rosen CA;
XX
DR   WPI; 1996-239509/24.
DR   P-PSDB; AAR95831.
XX
PT   Human ICE-LAP-3 and -4 DNA and protein - useful in the diagnosis
PT   and treatment of Alzheimer's disease, Parkinson's disease,
PT   rheumatoid arthritis, septic shock and head injury
XX
PS   Claim 1; Page 44; 67pp; English.
XX
CC   This sequence was derived from a cDNA library from human endometrial
CC   tumuor and may be used for the recombinant production of the ICE-LAP-
CC   4 protein e.g. in COS cells.  The encoded protein may be used
CC   therapeutically, e.g. as an antitumor or antiviral agent and to
CC   control embryonic development and tissue homeostasis.  The protein
CC   can also be used to treat immunosuppression disorders, such as AIDS,
CC   by targeting virus infected cells for cell death.  The DNA may find
CC   use in gene therapy applications.
XX
SQ   Sequence 1159 BP; 329 A; 220 C; 276 G; 334 T; 0 other;
```

```
Query Match             16.7%;  Score 229;  DB 17;  Length 1159;
Best Local Similarity   58.6%;  Pred. No. 3.5e-60;
Matches  439;  Conservative    0;  Mismatches  300;  Indels   10;  Gaps    2;

Qy    517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576

Db    331 tataaaatggattatcctgagatgggtttatgtataataattaataataagaattttcat 390

Qy    577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636

Db    391 aaaagcactggaatgacatctcggtctggtacagatgtcgatgcagcaaacctcagggaa 450

Qy    637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696

Db    451 acattcagaaacttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaaatt 510

Qy    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756

Db    511 gtggaattgatgcgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtttgt 570

Qy    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816

Db    571 gtgcttctgagccatggtgaagaaggaataatttttggaacaaatggacctgttgacctg 630

Qy    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876

Db    631 aaaaaaataacaaacttttcagaggggatcgttgtagaagtctaactggaaaacccaaa 690

Qy    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936

Db    691 cttttcattattcaggcctgccgtggtacagaactggactgtggcattgagacagacagt 750

Qy    937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996

Db    751 gg---------tgttgatgatgacatggcgtgtcataaaataccagtggaggccgacttc 801

Qy    997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056

Db    802 ttgtatgcatactccacagcacctggttattattcttggcgaaattcaaaggatggctcc 861

Qy   1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116

Db    862 tggttcatccagtcgctttgtgccatgctgaaacagtatgccgacaagcttgaatttatg 921

Qy   1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176

Db    922 cacattct-tacccgggttaaccgaaaggtggcaacagaatttgagtccttttccttga 980

Qy   1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
```

```
Db    981 cgctacttttcatgcaaagaaacagattccatgtattgtttccatgctcacaaaagaact 1040

Qy   1237 CTACTTCAGTCAATAGCCATATCAGGGGT 1265
Db   1041 ctatttttatcactaaagaaatggttggt 1069


RESULT  13
AAS10503
ID   AAS10503 standard; DNA; 2635 BP.
XX
AC   AAS10503;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Human caspase 3 DNA sequence.
XX
KW   Human; caspase 3; apoptosis; hyperproliferative disorder; hepatitis;
KW   viral infection; haematopoietic disorder; autoimmune disorder;
KW   atherosclerosis; neurological disorder; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             225..1058
FT                   /*tag= a
FT                   /product= "Caspase 3"
XX
PN   WO200153310-A1.
XX
PD   26-JUL-2001.
XX
PF   11-JAN-2001; 2001WO-US00888.
XX
PR   18-JAN-2000; 2000US-0484617.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Zhang H,  Cowsert LM;
XX
DR   WPI; 2001-442252/47.
DR   P-PSDB; AAU05394.
XX
PT   New antisense compound to inhibit caspase 3 is useful for treating
PT   hepatitis and atherosclerosis -
XX
PS   Example 13; Page 92-94; 127pp; English.
XX
CC   The present sequence encoding for human caspase 3 is described
CC   in an invention relating to novel antisense oligonucleotides
CC   (AAS10517-AAS10676) and methods of using these compounds for the
CC   modulation of caspase 3 expression. The caspase 3 antisense
CC   oligonucleotides specifically hybridise with and inhibit the
CC   expression of caspase 3. Antisense compounds targeted to caspase 3
CC   are useful to inhibit caspase 3 expression in cells or tissues and
CC   to modulate apoptosis. The caspase 3 antisense oligonucleotides are
CC   useful for treating disorders associated with expression of caspase 3.
CC   Such disorders include hyperproliferative disorders (e.g. cancer),
CC   viral infections (e.g. hepatitis), haematopoietic disorders, autoimmune
CC   disorders, atherosclerosis and neurological disorders (e.g. Alzheimer's
CC   disease).
XX
SQ   Sequence 2635 BP; 815 A; 465 C; 569 G; 786 T; 0 other;


  Query Match             16.7%;  Score 229;  DB 22;  Length 2635;
  Best Local Similarity   58.6%;  Pred. No. 5.9e-60;
  Matches  439;  Conservative    0;  Mismatches  300;  Indels   10;  Gaps    2;

Qy    517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
Db    333 tataaaatggattatcctgagatgggtttatgtataataattaataataagaatttttcat 392

Qy    577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
Db    393 aaaagcactggaatgacatctcggtctggtacagatgtcgatgcagcaaacctcagggaa 452

Qy    637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
Db    453 acattcagaaacttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaaatt 512

Qy    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
Db    513 gtggaattgatgcgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtttgt 572

Qy    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
Db    573 gtgcttctgagccatggtgaagaaggaataatttttggaacaaatggacctgttgacctg 632

Qy    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
Db    633 aaaaaaataacaaacttttcagaggggatcgttgtagaagtctaactggaaaacccaaa 692

Qy    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
Db    693 cttttcattattcaggcctgccgtggtacagaactggactgtggcattgagacagacagt 752

Qy    937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
Db    753 gg---------tgttgatgatgacatggcgtgtcataaaatacagtggatgccgacttc 803

Qy    997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
Db    804 ttgtatgcatactccacagcacctggttattattcttggcgaaattcaaaggatggctcc 863

Qy   1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
Db    864 tggttcatccagtcgctttgtgccatgctgaaacagtatgccgacaagcttgaatttatg 923

Qy   1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
Db    924 cacattct-tacccgggttaaccgaaaggtggcaacagaatttgagtccttttccctttga 982

Qy   1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
Db    983 cgctacttttcatgcaaagaaacagattccatgtattgtttccatgctcacaaaagaact 1042

Qy   1237 CTACTTCAGTCAATAGCCATATCAGGGGT 1265
Db   1043 ctatttttatcactaaagaaatggttggt 1071


RESULT  14
AAD03910
ID   AAD03910 standard; cDNA; 2635 BP.
XX
AC   AAD03910;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Human caspase-3 cDNA.
XX
KW   Human; caspase-3; CPP32; yama; apopain;
KW   cysteine protease; apoptosis; caspase expression cassette; metastasis;
KW   tumour; cathepsin B; urokinase; proliferation; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             225..1058
FT                   /*tag= a
FT                   /product= "Human caspase-3"
XX
PN   WO200129232-A2.
XX
```

```
PD   26-APR-2001.
XX
PF   19-OCT-2000; 2000WO-US28941.
XX
PR   20-OCT-1999;   99US-0160559.
PR   14-AUG-2000; 2000US-0225564.
XX
PA   (SCIO-) SCIOS INC.
XX
PI   Cordell B,  Li Y;
XX
DR   WPI; 2001-290920/30.
DR   P-PSDB; AAE00600.
XX
PT   Novel fusion polypeptide comprising first and second caspase subunit
PT   separated by cleavage site not associated in nature with caspase
PT   subunit, useful for cloning gene encoding enzymes involved in
PT   proteolytic cleavage  -
XX
PS   Example 1; Fig 5; 116pp; English.
XX
CC   The present sequence is a cDNA encoding human Caspase-3 also known as
CC   CPP32, yama and apopain. Caspases are a family of cysteine proteases,
CC   that participate in the initiation and execution of apoptosis. Caspases
CC   exist as pro-enzymes, activated by cleavage into a large and small
CC   subunit, occurring after specific aspartic acid residues within the
CC   pro-enzyme sequence. The present invention relates to a method for
CC   functional cloning of genes encoding proteins or enzymes involved in
CC   proteolytic cleavage. The invention is based on the use of caspase
CC   expression cassettes comprising the coding sequence of a proteolytic
CC   cleavage site flanked by sequences encoding two caspase subunits. A
CC   fusion polypeptide comprising a first and a second caspase subunit,
CC   separated by a cleavage site not associated in nature, is useful for
CC   cloning gene encoding enzymes involved in proteolytic cleavage. An
CC   expression cassette containing fusion polypeptide is used to identify a
CC   mutant cell line deficient in an enzyme of interest and is also useful
CC   for diagnosis and suppression of proliferation or metastases of a tumour
CC   cell characterised by overexpression of a polypeptide (e.g. Cathepsin B
CC   or urokinase, selectively expressed in the tumour cells). DNA encoding
CC   fusion polypeptide is used in gene therapy.
XX
SQ   Sequence 2635 BP; 815 A; 465 C; 569 G; 786 T; 0 other;


  Query Match             16.7%;  Score 229;  DB 22;  Length 2635;
  Best Local Similarity   58.6%;  Pred. No. 5.9e-60;
  Matches  439;  Conservative    0;  Mismatches  300;  Indels   10;  Gaps    2;

Qy     517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
           || || ||| ||| |    || ||||   ||| || ||||| || || ||||| ||| ||
Db     333 tataaaatggattatcctgagatgggtttatgtataataattaataataagaatttttcat 392

Qy     577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
           |||   || || |||     |||    || |||||    ||||| |     |||    |
Db     393 aaaagcactggaatgacatctcggtctggtacagatgtcgatgcagcaaacctcagggaa 452

Qy     637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
              ||| ||| | ||   | ||| || |      | |||||     |  |||   | ||
Db     453 acattcagaaacttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaaatt 512

Qy     697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
             ||  || |     | | |||| |||| || || |  ||      | | || |  ||
Db     513 gtggaattgatgcgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtttgt 572

Qy     757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
           | ||  | |||||||| ||||||    |||||| ||| | | ||||          |
Db     573 gtgcttctgagccatggtgaagaaggaataatttttggaacaaatggacctgttgacctg 632

Qy     817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
           ||  |  | |||  |   || || |||||| | || | ||  ||    |  |||||||||
Db     633 aaaaaaataacaaacttttcagaggggatcgttgtagaagtctaactggaaaacccaaa 692

Qy     877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
```

```
           || ||| | |||||||| ||||| || || || || ||   ||||||  || | |||
Db     693 cttttcattattcaggcctgccgtggtacagaactggactgtggcattgagacagacagt 752

Qy     937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
           ||          ||     |||| |   |  |  | || || |||||||| || ||||||
Db     753 gg---------tgttgatgatgacatggcgtgtcataaaataccagtggatgccgacttc 803

Qy     997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
             | |  || || ||||| |  || || ||||| || ||| | |   ||      ||||||
Db     804 ttgtatgcatactccacagcacctggttattattcttggcgaaattcaaaggatggctcc 863

Qy     1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
            |||||  | ||  | || ||  |||| ||| |  || | |   |  | || ||| | |||
Db     864 tggttcatccagtcgctttgtgccatgctgaaacagtatgccgacaagcttgaatttatg 923

Qy     1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
            || || ||  ||| |||| ||     |  || || |   | |||||||    ||   |||
Db     924 cacattct-tacccgggttaaccgaaaggtggcaacagaatttgagtccttttcctttga 982

Qy     1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
            | |    || |||  |||| |||| || ||| | || |||||||||| || |||||
Db     983 cgctacttttcatgcaaagaaacagattccatgtattgtttccatgctcacaaaagaact 1042

Qy     1237 CTACTTCAGTCAATAGCCATATCAGGGGT 1265
            ||| ||   ||| ||   | ||    |||
Db     1043 ctatttttatcactaaagaaatggttggt 1071


RESULT  15
AAD13075
ID   AAD13075 standard; cDNA; 834 BP.
XX
AC   AAD13075;
XX
DT   16-OCT-2001  (first entry)
XX
DE   Mammalia caspase-3 reporter moiety cDNA.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance;
KW   caspase-3; ss.
XX
OS   Mammalia.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US03791.
XX
PR   04-FEB-2000; 2000US-0498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by
PT   -NH-ubiquitin protein endoproteases, comprising two copies of the
PT   domain  -
XX
PS   Example 4; Page 131; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method
```

CC is also useful for developing novel assays for a wide range of
CC post-translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is mammalia caspase-3 cDNA which is used as a reporter moiety in
CC the exemplification of the invention.
XX
SQ Sequence 834 BP; 274 A; 147 C; 179 G; 234 T; 0 other;

Query Match 16.6%; Score 227.8; DB 22; Length 834;
Best Local Similarity 58.9%; Pred. No. 6.6e-60;
Matches 433; Conservative 0; Mismatches 292; Indels 10; Gaps 2;

```
Qy    517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
          || || ||| ||| |    || ||||  ||| || ||||| || || ||||| ||| ||
Db    109 tataaaatggattatcctgagatgggtttatgtataataattaataataagaattttcat 168

Qy    577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
          |||   || || |||     |||    || |||||    ||||| |     |||    |
Db    169 aaaagcactggaatgacatctcggtctggtacagatgtcgatgcagcaaacctcagggaa 228

Qy    637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
             ||| ||| | ||   | ||| || |      | |||||     |  |||   | ||
Db    229 acattcagaaacttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaaatt 288

Qy    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
             ||  || |     | | |||| |||| || || |  ||      | | || |  ||
Db    289 gtggaattgatgcgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtttgt 348

Qy    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
           | ||  | |||||||| ||||||    |||||| ||| | | ||||          |
Db    349 gtgcttctgagccatggtgaagaaggaataatttttggaacaaatggacctgttgacctg 408

Qy    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
          ||  |  | |||  |   || || |||||| | || | ||  ||    |  |||||||||
Db    409 aaaaaaataacaaacttttcagaggggatcgttgtagaagtctaactggaaaacccaaa 468

Qy    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
          || ||| | |||||||| ||||| || || || || ||   ||||||  || | |||
Db    469 cttttcattattcaggcctgccgtggtacagaactggactgtggcattgagacagacagt 528

Qy    937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
          ||          ||     ||||  |   |  |  | || || |||||||| || ||||||
Db    529 gg---------tgttgatgatgacatggcgtgtcataaaataccagtggaggccgacttc 579

Qy    997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
           | |  || || ||||| |  || || ||||| || ||| | |   ||      ||||||
Db    580 ttgtatgcatactccacagcacctggttattattcttggcgaaattcaaaggatggctcc 639

Qy   1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
          |||||  | ||  | || ||  |||| ||| |  || | |   |  | || ||| | |||
Db    640 tggttcatccagtcgctttgtgccatgctgaaacagtatgccgacaagcttgaatttatg 699

Qy   1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
          || || ||  ||| |||| ||    |  || || | |  | ||||||||    ||  |||
Db    700 cacattct-tacccgggttaaccgaaaggtggcaacagaatttgagtccttttcctttga 758

Qy   1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
          | |    || ||||  ||||| ||||| || ||| | || ||||||||||| || |||||
Db    759 cgctacttttcatgcaaagaaacagattccatgtattgtttccatgctcacaaaagaact 818

Qy   1237 CTACTTCAGTCAATA 1251
          ||| ||   ||| ||
Db    819 ctatttttatcacta 833
```

Search completed: August 16, 2002, 00:05:10
Job time: 6216 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2002, 22:16:34 ; Search time 74.75 Seconds
(without alignments)
4505.198 Million cell updates/sec

Title: US-09-613-508-1
Perfect score: 1371
Sequence: 1 GCACGAGAAACTTTGCTGTG.....TAATTTCAGGAAAATGTATT 1371

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1371	100.0	1371	3	US-08-462-969B-1	Sequence 1, Appli
2	232.8	17.0	1297	4	US-09-484-617-10	Sequence 10, Appl
3	229	16.7	1158	3	US-08-591-605-1	Sequence 1, Appli
4	229	16.7	1159	3	US-08-462-969B-3	Sequence 3, Appli
5	229	16.7	2635	4	US-09-484-617-3	Sequence 3, Appli
6	227.4	16.6	1001	3	US-08-964-308-5	Sequence 5, Appli
7	227.4	16.6	1001	3	US-08-964-313-5	Sequence 5, Appli
8	227.4	16.6	1001	4	US-09-069-138-5	Sequence 5, Appli
9	225.8	16.5	1001	3	US-08-964-308-9	Sequence 9, Appli
10	225.8	16.5	1001	3	US-08-964-313-9	Sequence 9, Appli
11	225.8	16.5	1001	4	US-09-069-138-9	Sequence 9, Appli
12	225.2	16.4	1177	2	US-08-890-542A-1	Sequence 1, Appli
13	225.2	16.4	1177	2	US-08-890-542A-3	Sequence 3, Appli
14	139	10.1	2658	2	US-08-773-608A-1	Sequence 1, Appli
15	105.6	7.7	1183	3	US-08-522-813-1	Sequence 1, Appli
16	105.6	7.7	1545	1	US-08-446-925-4	Sequence 4, Appli
17	105.6	7.7	1545	2	US-09-146-331-4	Sequence 4, Appli
18	105.6	7.7	1545	2	US-08-896-885-4	Sequence 4, Appli
19	105.6	7.7	1545	4	US-09-375-256-4	Sequence 4, Appli
20	97.6	7.1	1313	1	US-08-446-925-6	Sequence 6, Appli
21	97.6	7.1	1313	2	US-09-146-331-6	Sequence 6, Appli
22	97.6	7.1	1313	2	US-08-896-885-6	Sequence 6, Appli
23	97.6	7.1	1313	4	US-09-375-256-6	Sequence 6, Appli
24	85.8	6.3	1489	4	US-09-487-445-10	Sequence 10, Appl
25	82.2	6.0	2619	5	PCT-US96-10521-17	Sequence 17, Appl
26	82.2	6.0	2887	5	PCT-US96-10521-14	Sequence 14, Appl
27	80.6	5.9	870	4	US-09-360-017-2	Sequence 2, Appli
28	80.6	5.9	1286	2	US-08-807-200-3	Sequence 3, Appli
29	80.6	5.9	1286	4	US-09-001-777-3	Sequence 3, Appli
30	80.6	5.9	1330	2	US-08-807-200-1	Sequence 1, Appli
31	80.6	5.9	1330	4	US-09-001-777-1	Sequence 1, Appli
32	80.6	5.9	1355	2	US-08-618-408B-3	Sequence 3, Appli
33	80.6	5.9	1440	2	US-08-807-200-11	Sequence 11, Appl
34	80.6	5.9	1440	3	US-08-852-782-2	Sequence 2, Appli
35	80.6	5.9	1440	4	US-09-001-777-11	Sequence 11, Appl
36	80.6	5.9	1644	3	US-08-852-782-1	Sequence 1, Appli
37	80.6	5.9	1883	1	US-08-665-220-3	Sequence 3, Appli
38	80.6	5.9	1883	4	US-09-487-445-3	Sequence 3, Appli
39	80.6	5.9	1883	4	US-09-291-692-3	Sequence 3, Appli
40	80.6	5.9	2649	2	US-08-718-964-1	Sequence 1, Appli
41	80.6	5.9	2649	2	US-09-059-964A-1	Sequence 1, Appli
42	80.6	5.9	2649	2	US-08-842-341-1	Sequence 1, Appli
43	79	5.8	1443	5	PCT-US96-10521-33	Sequence 33, Appl
44	73.8	5.4	1700	1	US-08-665-220-1	Sequence 1, Appli
45	73.8	5.4	1700	4	US-09-291-692-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-08-462-969B-1
; Sequence 1, Application US/08462969B
; Patent No. 6087150
;  GENERAL INFORMATION:
;    APPLICANT:  He, Wei-Wu et al.
;    TITLE OF INVENTION:  Interleukin-1 Beta Converting Enzyme
;    TITLE OF INVENTION:  Like Apoptosis Protease 3 and 4
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Ave.
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,969B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/334,251
;      FILING DATE:  11-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF140P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1371 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-462-969B-1


  Query Match           100.0%;  Score 1371;  DB 3;  Length 1371;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1371;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCACGAGAAACTTTGCTGTGCGCGTTCTCCCGCGCGCGGGCTCAACTTTGTAGAGCGAGG 60
```

```
Db      1 GCACGAGAAACTTTGCTGTGCGCGTTCTCCCGCGCGCGGGCTCAACTTTGTAGAGCGAGG 60

Qy     61 GGCCAACTTGGCAGAGCGCGCGGCCAGCTTTGCAGAGAGCGCCCTCCAGGGACTATGCGT 120
Db     61 GGCCAACTTGGCAGAGCGCGCGGCCAGCTTTGCAGAGAGCGCCCTCCAGGGACTATGCGT 120

Qy    121 GCGGGGACACGGGTCGCTTTGGGCTCTTCCACCCCTGCGGAGCGCACTACCCCGAGCCAG 180
Db    121 GCGGGGACACGGGTCGCTTTGGGCTCTTCCACCCCTGCGGAGCGCACTACCCCGAGCCAG 180

Qy    181 GGGCGGTGCAAGCCCCGCCCGGCCCTACCCAGGGCGGCTCCTCCCTCCGCAGCGCCGAGA 240
Db    181 GGGCGGTGCAAGCCCCGCCCGGCCCTACCCAGGGCGGCTCCTCCCTCCGCAGCGCCGAGA 240

Qy    241 CTTTTAGTTTCGCTTTCGCTAAAGGGGCCCCAGACCCTTGCTGCGGAGCGACGGAGAGAG 300
Db    241 CTTTTAGTTTCGCTTTCGCTAAAGGGGCCCCAGACCCTTGCTGCGGAGCGACGGAGAGAG 300

Qy    301 ACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGGGCTGT 360
Db    301 ACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGGGCTGT 360

Qy    361 ATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGAC 420
Db    361 ATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGAC 420

Qy    421 CGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCC 480
Db    421 CGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCC 480

Qy    481 ATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTG 540
Db    481 ATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTG 540

Qy    541 GGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGA 600
Db    541 GGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGA 600

Qy    601 AACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGAC 660
Db    601 AACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGAC 660

Qy    661 GTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCT 720
Db    661 GTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCT 720

Qy    721 GAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAA 780
Db    721 GAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAA 780

Qy    781 AATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGG 840
Db    781 AATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGG 840

Qy    841 GGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGA 900
Db    841 GGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGA 900

Qy    901 GGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCT 960
Db    901 GGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCT 960

Qy    961 AATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCA 1020
Db    961 AATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCA 1020

Qy   1021 GGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCC 1080
Db   1021 GGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCC 1080

Qy   1081 ATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGA 1140
Db   1081 ATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGA 1140

Qy   1141 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCA 1200
Db   1141 CAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCA 1200

Qy   1201 GATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCA 1260
Db   1201 GATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCA 1260

Qy   1261 GGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGC 1320
Db   1261 GGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGC 1320

Qy   1321 TCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
Db   1321 TCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
```

RESULT 2
US-09-484-617-10
; Sequence 10, Application US/09484617
; Patent No. 6303374
; GENERAL INFORMATION:
; APPLICANT: Hong Zhang
; APPLICANT: Lex M. Cowsert
; TITLE OF INVENTION: ANTISENSE MODULATION OF CASPASE 3 EXPRESSION
; FILE REFERENCE: RTS-0103
; CURRENT APPLICATION NUMBER: US/09/484,617
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 176
; SEQ ID NO 10
; LENGTH: 1297
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (68)...(901)
US-09-484-617-10

Query Match 17.0%; Score 232.8; DB 4; Length 1297;
Best Local Similarity 59.0%; Pred. No. 3e-64;
Matches 441; Conservative 0; Mismatches 297; Indels 10; Gaps 2;

```
Qy    517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
Db    176 tacaaaatggattatcctgaaatgggcatatgcataataattaataataagaaacttccat 235

Qy    577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
Db    236 aagagcactggaatgtcatctcgctctggtacggatgtggacgcagccaacctcagagag 295

Qy    637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
Db    296 acattcatgggcctgaaataccaagtcaggaataaaaatgatcttactcgtgaagacatt 355

Qy    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
Db    356 ttggaattaatggatagtgtttctaaggaagatcatagcaaaaggagcagctttgtgtgt 415

Qy    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
Db    416 gtgattctaagccatggtgatgaaggggtcatttatgggacaaatgggcctgttgaactg 475

Qy    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
Db    476 aaaaagttgactagcttcttcagaggcgactactgccggagtctgactggaaagccgaaa 535

Qy    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
Db    536 ctcttcatcattcaggcctgccggggtacggagctggactgtggcattgagacagacagt 595
```

```
Qy    937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
          ||| |     | || |  |         || |  | ||||| || ||||| |||||||||
Db    596 gggactgatgaggagatgg---------cttgccagaagataccggtggaggctgacttc 646

Qy    997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
          || |  || || || || |  || || || || || ||||| |   ||      || ||
Db    647 ctgtatgcttactctacagcacctggttactattcctggagaaattcaaaggacgggtcg 706

Qy   1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
          |||||  | ||  |||| |||  ||| ||| ||  | |||   |  | || ||| | |||
Db    707 tggttcatccagtccctttgcagcatgctgaagctgtacgcgcacaagctagaatttatg 766

Qy   1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
          || || || |||  | || ||    |  || || | | | || |||||    ||    ||
Db    767 cacattct-cactcgcgttaacaggaaggtggcaacggaattcgagtccttctccctgga 825

Qy   1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
          | |    ||||| |  ||||| |||||||| ||| | || ||||||||||| || |||||
Db    826 ctccactttccacgcaaagaaacagatcccgtgtattgtgtccatgctcacgaaagaact 885

Qy   1237 CTACTTCAGTCAATAGCCATATCAGGGG 1264
           |||||   ||| |||    || |  ||
Db    886 gtacttttatcactagaggaatgattgg 913
```

```
RESULT   3
US-08-591-605-1
; Sequence 1, Application US/08591605
; Patent No. 6060238
;  GENERAL INFORMATION:
;    APPLICANT:  Dixit, Vishva M.
;    TITLE OF INVENTION:  METHOD AND COMPOSITION FOR REGULATING
;    TITLE OF INVENTION:  APOPTOSIS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 PAGE MILL ROAD
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/591,605
;      FILING DATE:  09-FEB-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KONSKI, ANTOINETTE F.
;      REGISTRATION NUMBER:  34,202
;      REFERENCE/DOCKET NUMBER:  20344-21036.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  706141 MRSNFOERS SFO
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1158 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  224..1054
;      OTHER INFORMATION:  /product= "Yama peptide sequence"
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  704..718
;      OTHER INFORMATION:  /note= "region coding for
;      OTHER INFORMATION:  pentapeptide, wherein cysteine is catalytic"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "atg")
;      OTHER INFORMATION:  /note= "Mutation of catalytic cysteine to methionine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gcc")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gcg")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gct")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gca")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
US-08-591-605-1
```

```
  Query Match             16.7%;  Score 229;  DB 3;  Length 1158;
  Best Local Similarity   58.6%;  Pred. No. 4.5e-63;
  Matches  439;  Conservative    0;  Mismatches  300;  Indels   10;  Gaps    2;

Qy    517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
          || || ||| ||| |    || ||||   ||| || ||||| || || ||||| ||| ||
Db    332 TATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCAT 391

Qy    577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
          |||   || || |||     |||    || |||||   ||||| |    |||     |
Db    392 AAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAA 451

Qy    637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
            ||| ||| | ||   | ||| || |     | |||||     |  |||   | ||
Db    452 ACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATT 511

Qy    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
          ||  || |     | | |||| |||| || || |  ||      | | || |  ||
Db    512 GTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGT 571

Qy    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
          | ||  | |||||||| ||||||    |||||| ||| | | ||||         |
Db    572 GTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTG 631

Qy    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
          ||  |  | |||  |   || || |||||| | || | ||  ||    |  ||||||||||
Db    632 AAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAA 691

Qy    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
          || ||| | |||||||| ||||| || || || || ||   ||||||  || | |||
Db    692 CTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGT 751

Qy    937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
          ||          ||      |||| |   |  |  | || || |||||||| || ||||||
Db    752 GG---------TGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTC 802

Qy    997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
          | |  || || ||||| |  || || ||||| || ||| | |   ||      ||||||
Db    803 TTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCC 862

Qy   1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
          |||||  | ||  | || ||  |||| ||| |  || | |   |  | || ||| | |||
Db    863 TGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATG 922

Qy   1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
          || || ||  ||| |||| ||     |  || || |   | ||||||||     ||  |||
Db    923 CACATTCT-TACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGA 981

Qy   1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
          | |    || ||||  ||||| ||||| || ||| | || ||||||||||| || |||||
```

```
Db    982 CGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACT 1041

Qy   1237 CTACTTCAGTCAATAGCCATATCAGGGGT 1265
          ||| ||   ||| ||   | ||    |||
Db   1042 CTATTTTTATCACTAAAGAAATGGTTGGT 1070


RESULT   4
US-08-462-969B-3
; Sequence 3, Application US/08462969B
; Patent No. 6087150
;  GENERAL INFORMATION:
;    APPLICANT:  He, Wei-Wu et al.
;    TITLE OF INVENTION:  Interleukin-1 Beta Converting Enzyme
;    TITLE OF INVENTION:  Like Apoptosis Protease 3 and 4
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Ave.
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,969B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/334,251
;      FILING DATE:  11-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF140P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1159 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-462-969B-3


  Query Match             16.7%;  Score 229;  DB 3;  Length 1159;
  Best Local Similarity   58.6%;  Pred. No. 4.5e-63;
  Matches  439;  Conservative    0;  Mismatches  300;  Indels   10;  Gaps    2;

Qy    517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
          || || ||| ||| |    || ||||   ||| || ||||| || || ||||| ||| ||
Db    331 TATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCAT 390

Qy    577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
          |||   || || |||     |||     || |||||    ||||| |     |||    |
Db    391 AAAAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAA 450

Qy    637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
            ||| ||| | ||   | ||| || |       | |||||     |  |||   | ||
Db    451 ACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATT 510

Qy    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
            || || |     | | |||| |||| || || |  ||       | | || |  ||
Db    511 GTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGT 570

Qy    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
          | ||  | |||||||| ||||||    |||||| ||| | | ||||           |
Db    571 GTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTG 630

Qy    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
          ||  |  | |||  |    || || |||||| | || | ||  ||     |  |||||||||
Db    631 AAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAA 690

Qy    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
          || ||| | |||||||| ||||| || || || || ||   ||||||  || | |||
Db    691 CTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGT 750

Qy    937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
          ||          ||     ||||  |   |  |  | || || |||||||| || ||||||
Db    751 GG---------TGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTC 801

Qy    997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
          | |  || || ||||| |  || || ||||| || ||| | |   ||      ||||||
Db    802 TTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCC 861

Qy   1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
          |||||  | ||  | || ||  |||| ||| |  || | |   |  | || ||| | |||
Db    862 TGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATG 921

Qy   1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
          || || ||  ||| |||| ||    |  || || |   | ||||||||    ||  |||
Db    922 CACATTCT-TACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGA 980

Qy   1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
          | |    || ||||  ||||| ||||| || ||| | || |||||||||| || |||||
Db    981 CGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACT 1040

Qy   1237 CTACTTCAGTCAATAGCCATATCAGGGGT 1265
          ||| ||   ||| ||   | ||    |||
Db   1041 CTATTTTTATCACTAAAGAAATGGTTGGT 1069


RESULT   5
US-09-484-617-3
; Sequence 3, Application US/09484617
; Patent No. 6303374
; GENERAL INFORMATION:
;  APPLICANT: Hong Zhang
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF CASPASE 3 EXPRESSION
;  FILE REFERENCE: RTS-0103
;  CURRENT APPLICATION NUMBER: US/09/484,617
;  CURRENT FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 176
; SEQ ID NO 3
;   LENGTH: 2635
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (225)...(1058)
US-09-484-617-3


  Query Match             16.7%;  Score 229;  DB 4;  Length 2635;
  Best Local Similarity   58.6%;  Pred. No. 7.8e-63;
  Matches  439;  Conservative    0;  Mismatches  300;  Indels   10;  Gaps    2;

Qy    517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
          || || ||| ||| |    || ||||   ||| || ||||| || || ||||| ||| ||
Db    333 tataaaatggattatcctgagatgggtttatgtataataattaataataagaattttcat 392

Qy    577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
          |||   || || |||     |||    || |||||    ||||| |     |||    |
Db    393 aaaagcactggaatgacatctcggtctggtacagatgtcgatgcagcaaacctcagggaa 452

Qy    637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
            ||| ||| | ||   | ||| || |       | |||||     |  |||   | ||
Db    453 acattcagaaacttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaaatt 512
```

```
Qy    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
Db    513 gtggaattgatgcgtgatgtttctaaagaagatcacagcaaaaggagcagttttgtttgt 572

Qy    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
Db    573 gtgcttctgagccatggtgaagaaggaataatttttggaacaaatggacctgttgacctg 632

Qy    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
Db    633 aaaaaaataacaaacttttcagaggggatcgttgtagaagtctaactggaaaacccaaa 692

Qy    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
Db    693 cttttcattattcaggcctgccgtggtacagaactggactgtggcattgagacagacagt 752

Qy    937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
Db    753 gg---------tgttgatgatgacatggcgtgtcataaaataccagtggatgccgacttc 803

Qy    997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
Db    804 ttgtatgcatactccacagcacctggttattattcttggcgaaattcaaaggatggctcc 863

Qy   1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
Db    864 tggttcatccagtcgctttgtgccatgctgaaacagtatgccgacaagcttgaatttatg 923

Qy   1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
Db    924 cacattct-taccgggttaaccgaaaggtggcaacagaatttgagtcctttccctttga 982

Qy   1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
Db    983 cgctacttttcatgcaaagaaacagattccatgtattgtttccatgctcacaaaagaact 1042

Qy   1237 CTACTTCAGTCAATAGCCATATCAGGGGT 1265
Db   1043 ctatttttatcactaaagaaatggttggt 1071
```

RESULT 6
US-08-964-308-5
; Sequence 5, Application US/08964308
; Patent No. 6066715
; GENERAL INFORMATION:
; APPLICANT: DESMARAIS, SYLVIE
; APPLICANT: FRIESEN, RICHARD
; APPLICANT: ZAMBONI, ROBERT
; TITLE OF INVENTION: NEW LIGANDS FOR PHOSPHATASE
; TITLE OF INVENTION: BINDING ASSAY
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBERT J. NORTH - MERCK & CO., INC.
; STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Diskette
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/964,308
; FILING DATE: 04-NOV-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: NORTH, ROBERT J
; REGISTRATION NUMBER: 27,366
; REFERENCE/DOCKET NUMBER: 19840 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-7262
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-964-308-5

Query Match 16.6%; Score 227.4; DB 3; Length 1001;
Best Local Similarity 58.5%; Pred. No. 1.3e-62;
Matches 438; Conservative 0; Mismatches 301; Indels 10; Gaps 2;

```
Qy    517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
Db    172 TATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCAT 231

Qy    577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
Db    232 AAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAA 291

Qy    637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
Db    292 ACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATT 351

Qy    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
Db    352 GTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGT 411

Qy    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
Db    412 GTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTG 471

Qy    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
Db    472 AAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAA 531

Qy    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
Db    532 CTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGT 591

Qy    937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
Db    592 GG---------TGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTC 642

Qy    997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
Db    643 TTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCC 702

Qy   1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
Db    703 TGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATG 762

Qy   1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
Db    763 CACATTCT-TACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGA 821

Qy   1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
Db    822 CGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACT 881

Qy   1237 CTACTTCAGTCAATAGCCATATCAGGGGT 1265
Db    882 CTATTTTTATCACTAAAGAAATGGTTGGT 910
```

RESULT 7
US-08-964-313-5

```
; Sequence 5, Application US/08964313
; Patent No. 6114132
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  GRESSER, MICHAEL
;    APPLICANT:  KENNEDY, BRIAN
;    APPLICANT:  NICHOLSON, DONALD
;    APPLICANT:  RAMACHANDRAN, CHIDAMBARAN
;    APPLICANT:  SKOREY, KATHRYN
;    APPLICANT:  FORD-HUTCHINSON, ANTHONY
;    TITLE OF INVENTION:  PHOSPHATASE BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,313
;      FILING DATE:  04-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/030,408
;      FILING DATE:  04-NOV-1996
;      APPLICATION NUMBER:  PCT/CA97/00825
;      FILING DATE:  03 -NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19824Y
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-313-5


  Query Match           16.6%;  Score 227.4;  DB 3;  Length 1001;
  Best Local Similarity  58.5%;  Pred. No. 1.3e-62;
  Matches  438;  Conservative    0;  Mismatches  301;  Indels   10;  Gaps    2;

Qy     517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
           || || ||| ||| |    || ||||   ||| || ||||| || || ||||| ||| ||
Db     172 TATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCAT 231

Qy     577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
           ||    || || |||     |||    || |||||    ||||| |     |||    |
Db     232 AAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAA 291

Qy     637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
             ||| ||| | ||    | ||| || |      | |||||     |  |||   | ||
Db     292 ACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATT 351

Qy     697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
             ||  || |     | | |||| |||| || || |  ||       | | || |  ||
Db     352 GTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGT 411

Qy     757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
           | ||  | |||||||| ||||||    |||||| ||| | | ||||          |
Db     412 GTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTG 471

Qy     817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
           ||  |  | |||  |   || || |||||| | || | ||  ||     |  |||||||||
Db     472 AAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAA 531

Qy     877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
           || ||| | |||||||| ||||| || || || || ||   ||||||  || | |||
Db     532 CTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGT 591

Qy     937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
           ||          ||     |||| |   |  |  | || || |||||||| || ||||||
Db     592 GG---------TGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTC 642

Qy     997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
           | |  || || ||||| |  || || ||||| || ||| | |   ||      ||||||
Db     643 TTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCC 702

Qy     1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
           |||||  | ||  | ||  ||  |||| ||| |  || | |   |  | || ||| | |||
Db     703 TGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATG 762

Qy     1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
           || || ||  ||| |||| ||    |  || || |   | ||||||||    ||  |||
Db     763 CACATTCT-TACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGA 821

Qy     1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
           | |    || ||||  ||||| ||||| || ||| | || |||||||||| || |||||
Db     822 CGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACT 881

Qy     1237 CTACTTCAGTCAATAGCCATATCAGGGGT 1265
           |||  ||   ||| ||   | ||    |||
Db     882 CTATTTTTATCACTAAAGAAATGGTTGGT 910


RESULT   8
US-09-069-138-5
; Sequence 5, Application US/09069138
; Patent No. 6348572
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  DUFRESNE, CLAUDE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  LEBLANC, YVES
;    APPLICANT:  ROY, PATRICK
;    APPLICANT:  YOUNG, ROBERT N.
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/069,138
;      FILING DATE:  29-APR-1998
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19840YIA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
```

```
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-069-138-5


  Query Match           16.6%;  Score 227.4;  DB 4;  Length 1001;
  Best Local Similarity  58.5%;  Pred. No. 1.3e-62;
  Matches  438;  Conservative    0;  Mismatches  301;  Indels   10;  Gaps    2;

Qy    517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
          || || ||| ||| |    || ||||   ||| || ||||| || || ||||| ||| ||
Db    172 TATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCAT 231

Qy    577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
          ||    || || |||     |||    || |||||    ||||| |     |||    |
Db    232 AAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAA 291

Qy    637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
             ||| ||| | ||   | ||| || |     | |||||     |  |||   | ||
Db    292 ACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATT 351

Qy    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
             ||  || |     | | |||| |||| || || |  ||       | | || |  ||
Db    352 GTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGT 411

Qy    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
            | ||  | |||||||| ||||||    |||||| ||| | | ||||           |
Db    412 GTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTG 471

Qy    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
          ||  |  | |||  |   || || ||||||  | ||  | ||  ||    |  |||||||||
Db    472 AAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAA 531

Qy    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
          || ||| | |||||||| ||||| || || || || ||   ||||||  || | |||
Db    532 CTTTTCATTATTCAGGCCTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGT 591

Qy    937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
          ||         ||     |||| |   |  | | || || |||||||| || ||||||
Db    592 GG---------TGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTC 642

Qy    997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
          | |  || || ||||| |  || || ||||| || ||| | |   ||       ||||||
Db    643 TTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCC 702

Qy   1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
          |||||  | ||  | || ||  |||| ||| |  || | |   |  | || ||| | |||
Db    703 TGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATG 762

Qy   1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
          || || ||  ||| |||| ||     |  || || |   | ||||||||     ||   |||
Db    763 CACATTCT-TACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGA 821

Qy   1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
          | |    || ||||  |||| |||| || ||| | || ||||||||||| || |||||
Db    822 CGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACT 881

Qy   1237 CTACTTCAGTCAATAGCCATATCAGGGGT 1265
          ||| ||   ||| ||   | ||   |||
Db    882 CTATTTTTATCACTAAAGAAATGGTTGGT 910


RESULT   9
US-08-964-308-9
; Sequence 9, Application US/08964308
; Patent No. 6066715
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBERT J. NORTH - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,308
;      FILING DATE:  04-NOV-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  NORTH, ROBERT J
;      REGISTRATION NUMBER:  27,366
;      REFERENCE/DOCKET NUMBER:  19840 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-7262
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-308-9


  Query Match           16.5%;  Score 225.8;  DB 3;  Length 1001;
  Best Local Similarity  58.3%;  Pred. No. 4.3e-62;
  Matches  437;  Conservative    0;  Mismatches  302;  Indels   10;  Gaps    2;

Qy    517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
          || || ||| ||| |    || ||||   ||| || ||||| || || ||||| ||| ||
Db    172 TATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCAT 231

Qy    577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
          ||    || || |||     |||    || |||||    ||||| |     |||    |
Db    232 AAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAA 291

Qy    637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
             ||| ||| | ||   | ||| || |     | |||||     |  |||   | ||
Db    292 ACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATT 351

Qy    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
             ||  || |     | | |||| |||| || || |  ||       | | || |  ||
Db    352 GTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGT 411

Qy    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
            | ||  | |||||||| ||||||    |||||| ||| | | ||||           |
Db    412 GTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTG 471

Qy    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
          ||  |  | |||  |   || || ||||||  | ||  | ||  ||    |  |||||||||
Db    472 AAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAA 531

Qy    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
          || ||| | |||||||| ||||| || || || || ||   ||||||  || | |||
Db    532 CTTTTCATTATTCAGGCCTCCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGT 591
```

```
Qy     937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
           ||          ||      |||| |    |  | | || || |||||||| || ||||||
Db     592 GG---------TGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTC 642

Qy     997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
           | |  || || ||||| |  || || ||||| || ||| | |   ||       ||||||
Db     643 TTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCC 702

Qy    1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
           |||||  | ||  | || ||  |||| ||| |  || | |    |  | || ||| | |||
Db     703 TGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATG 762

Qy    1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
           || || ||  ||| |||| ||    |  || || |  | ||||||||    ||   |||
Db     763 CACATTCT-TACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGA 821

Qy    1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
           | |    || ||||  ||||| ||||| || ||| | || ||||||||||| || |||||
Db     822 CGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACT 881

Qy    1237 CTACTTCAGTCAATAGCCATATCAGGGGT 1265
           ||| ||   ||| ||   | ||    |||
Db     882 CTATTTTTATCACTAAAGAAATGGTTGGT 910


RESULT  10
US-08-964-313-9
; Sequence 9, Application US/08964313
; Patent No. 6114132
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  GRESSER, MICHAEL
;    APPLICANT:  KENNEDY, BRIAN
;    APPLICANT:  NICHOLSON, DONALD
;    APPLICANT:  RAMACHANDRAN, CHIDAMBARAN
;    APPLICANT:  SKOREY, KATHRYN
;    APPLICANT:  FORD-HUTCHINSON, ANTHONY
;    TITLE OF INVENTION:  PHOSPHATASE BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,313
;      FILING DATE:  04-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/030,408
;      FILING DATE:  04-NOV-1996
;      APPLICATION NUMBER:  PCT/CA97/00825
;      FILING DATE:  03 -NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19824Y
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-313-9


  Query Match             16.5%;  Score 225.8;  DB 3;  Length 1001;
  Best Local Similarity   58.3%;  Pred. No. 4.3e-62;
  Matches  437;  Conservative    0;  Mismatches  302;  Indels   10;  Gaps    2;

Qy     517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
           || || ||| ||| |    || ||||   ||| || ||||| || || ||||| ||| ||
Db     172 TATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCAT 231

Qy     577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
           ||    || || |||     |||    || |||||   ||||| |    |||    |
Db     232 AAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAA 291

Qy     637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
           ||| ||| | ||   | ||| || |      | |||||     |  |||   | ||
Db     292 ACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATT 351

Qy     697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
           ||  || |     | | |||| |||| || || | ||      | | || |  ||
Db     352 GTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGT 411

Qy     757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
           | ||  | |||||||| |||||    |||||| ||| | | ||||             |
Db     412 GTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTG 471

Qy     817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
           ||  |  | |||  |   || || |||||| | || | ||  ||     |  |||||||||
Db     472 AAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAA 531

Qy     877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
           || ||| | |||||||| | ||| || || || || ||   ||||||  || | |||
Db     532 CTTTTCATTATTCAGGCCTCCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGT 591

Qy     937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
           ||          ||      |||| |    |  | | || || |||||||| || ||||||
Db     592 GG---------TGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTC 642

Qy     997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
           | |  || || ||||| |  || || ||||| || ||| | |   ||       ||||||
Db     643 TTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCC 702

Qy    1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
           |||||  | ||  | || ||  |||| ||| |  || | |    |  | || ||| | |||
Db     703 TGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATG 762

Qy    1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
           || || ||  ||| |||| ||    |  || || |  | ||||||||    ||   |||
Db     763 CACATTCT-TACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGA 821

Qy    1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
           | |    || ||||  ||||| ||||| || ||| | || ||||||||||| || |||||
Db     822 CGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACT 881

Qy    1237 CTACTTCAGTCAATAGCCATATCAGGGGT 1265
           ||| ||   ||| ||   | ||    |||
Db     882 CTATTTTTATCACTAAAGAAATGGTTGGT 910


RESULT  11
US-09-069-138-9
; Sequence 9, Application US/09069138
; Patent No. 6348572
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  DUFRESNE, CLAUDE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  LEBLANC, YVES
;    APPLICANT:  ROY, PATRICK
```

```
;    APPLICANT:  YOUNG, ROBERT N.
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/069,138
;      FILING DATE:  29-APR-1998
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19840YIA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-069-138-9

  Query Match           16.5%;  Score 225.8;  DB 4;  Length 1001;
  Best Local Similarity  58.3%;  Pred. No. 4.3e-62;
  Matches  437;  Conservative    0;  Mismatches  302;  Indels   10;  Gaps    2;

Qy     517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
           || || ||| ||| |    || ||||   ||| || ||||| || || ||||| ||| ||
Db     172 TATAAAATGGATTATCCTGAGATGGGTTTATGTATAATAATTAATAATAAGAATTTTCAT 231

Qy     577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
           ||    || || |||     |||    || |||||    ||||| |     |||    |
Db     232 AAGAGCACTGGAATGACATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAA 291

Qy     637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
              ||| ||| | ||   | ||| || |      | |||||     |  |||   | ||
Db     292 ACATTCAGAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAAATT 351

Qy     697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
             ||  || |     | | |||| |||| || || |  ||       | | || |  ||
Db     352 GTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAGCAGTTTTGTTTGT 411

Qy     757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
            | ||  | ||||||||| ||||||    |||||| ||| | | ||||           |
Db     412 GTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTGGAACAAATGGACCTGTTGACCTG 471

Qy     817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
           ||  |  | |||  |   || || |||||| | || | ||  ||    |  |||||||||
Db     472 AAAAAAATAACAAACTTTTTCAGAGGGGATCGTTGTAGAAGTCTAACTGGAAAACCCAAA 531

Qy     877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
           || ||| | |||||||| | ||| || || || || ||   ||||||  || | |||
Db     532 CTTTTCATTATTCAGGCCTCCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGT 591

Qy     937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
           ||         ||     ||||  |   |  |  | || || |||||||| || ||||||

Db     592 GG---------TGTTGATGATGACATGGCGTGTCATAAAATACCAGTGGAGGCCGACTTC 642

Qy     997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
             |  |  || || ||||| |  || || ||||| || ||| | |   ||       ||||||
Db     643 TTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTCAAAGGATGGCTCC 702

Qy    1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
           |||||  | ||  | || ||  |||| ||| |  || | |    |  | || ||| | |||
Db     703 TGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGTATGCCGACAAGCTTGAATTTATG 762

Qy    1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
           || || ||  ||| |||| ||    |  || || |   | ||||||||    ||   |||
Db     763 CACATTCT-TACCCGGGTTAACCGAAAGGTGGCAACAGAATTTGAGTCCTTTTCCTTTGA 821

Qy    1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
           | |    || |||| ||||| ||||| || ||| | || ||||||||||| || |||||
Db     822 CGCTACTTTTCATGCAAAGAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACT 881

Qy    1237 CTACTTCAGTCAATAGCCATATCAGGGGT 1265
           ||| ||   ||| ||   | ||    |||
Db     882 CTATTTTTATCACTAAAGAAATGGTTGGT 910


RESULT  12
US-08-890-542A-1
; Sequence 1, Application US/08890542A
; Patent No. 5840509
;  GENERAL INFORMATION:
;    APPLICANT:  Ni, Binhui
;    APPLICANT:  Paul, Steven M
;    APPLICANT:  Wu, Xin
;    TITLE OF INVENTION:  PROTEASE AND RELATED NUCLEIC ACID
;    TITLE OF INVENTION:  COMPOUNDS
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  United States of America
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/890,542A
;      FILING DATE:  09-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gaylo, Paul J.
;      REGISTRATION NUMBER:  36,808
;      REFERENCE/DOCKET NUMBER:  X-10704
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (317) 276-0756
;      TELEFAX:  (317) 276-3861
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1177 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  76..906
US-08-890-542A-1

  Query Match           16.4%;  Score 225.2;  DB 2;  Length 1177;
  Best Local Similarity  58.3%;  Pred. No. 7.5e-62;
  Matches  437;  Conservative    0;  Mismatches  303;  Indels   10;  Gaps    2;
```

```
Qy    517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
          ||||| ||| |||      |  ||||   || || ||||| || || ||||||||  ||
Db    184 TACAAAATGGATTACCCTGAAATGGGCTTGTGTATAATAATTAATAATAAGAACTTCCAT 243

Qy    577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
          |||   || || |||   | ||| || || || ||    |||||  |     |||    ||
Db    244 AAAAGCACTGGAATGTCAGCTCGCAATGGTACCGATGTCGATGCAGCTAACCTCAGAGAG 303

Qy    637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
            |||     ||||   |  || || |      | ||||||    || |||   | ||
Db    304 ACATTCATGGCCCTGAAATACGAAGTCAGGAATAAAAATGACCTTACTCGTGAAGAAATT 363

Qy    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
            ||  || |  | |  | |||| | || || || | ||      | | || |  ||
Db    364 ATGGAATTGATGGATAGTGTTTCTAAGGAAGATCACAGCAAAAGGAGCAGTTTTGTGTGT 423

Qy    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
            |  |  |||| |||||||| |||   |||||||  ||| |   | ||       |
Db    424 GTGATTCTAAGTCATGGAGATGAAGGAGTAATTTTTGGAACGAACGGACCTGTGGACCTG 483

Qy    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
          ||  |  | ||      ||| || || ||    |||   |  ||    |  || || |||
Db    484 AAAAAACTAACTAGTTTCTTCAGAGGCGACTACTGCCGGAGTCTGACTGGAAAGCCGAAA 543

Qy    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
          |||||| |||||||||| ||||||||  || ||||| ||    ||  ||  || | |||
Db    544 CTCTTCATCATTCAGGCCTGCCGAGGTACAGAGCTGGACTGCGGTATTGAGACAGACAGT 603

Qy    937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
          ||  |   |      |||| ||   |  |  | ||||| |||||||  || ||||||
Db    604 GGAACTGAC---------GATGATATGGCATGCCAGAAGATACCAGTGGGGGCCGACTTC 654

Qy    997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
          || |  || || || || |  || || || || || || ||||| |   || |    || ||
Db    655 CTGTATGCTTACTCTACCGCACCCGGTTACTATTCCTGGAGAAATTCAAGGGACGGGTCA 714

Qy   1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
          |||||  | ||  | || ||| |||| ||| |   | |||   |  | |||||| ||||
Db    715 TGGTTCATCCAGTCACTTTGCGCCATGCTGAAACTGTACGCGCACAAGCTGGAATTCATG 774

Qy   1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
          || ||||| |||  | || ||    |  || |||| | | |||||||    ||    ||
Db    775 CACATCCT-CACTCGTGTTAACCGGAAGGTGGCCATGGAATTTGAGTCCTTCTCCCTGGA 833

Qy   1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
          | |    ||||| |  ||||| |||||||| ||| | || || |||||||| || |||||
Db    834 CGCCACTTTCCACGCAAAGAAACAGATCCCGTGTATTGTGTCAATGCTCACAAAAGAACT 893

Qy   1237 CTACTTCAGTCAATAGCCATATCAGGGGTA 1266
            |||||   ||| ||     || |  || |
Db    894 GTACTTTTATCACTAAAGGAATGACTGGGA 923
```

```
RESULT 13
US-08-890-542A-3
; Sequence 3, Application US/08890542A
; Patent No. 5840509
;  GENERAL INFORMATION:
;    APPLICANT:  Ni, Binhui
;    APPLICANT:  Paul, Steven M
;    APPLICANT:  Wu, Xin
;    TITLE OF INVENTION:  PROTEASE AND RELATED NUCLEIC ACID
;    TITLE OF INVENTION:  COMPOUNDS
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  United States of America
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/890,542A
;      FILING DATE:  09-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gaylo, Paul J.
;      REGISTRATION NUMBER:  36,808
;      REFERENCE/DOCKET NUMBER:  X-10704
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (317) 276-0756
;      TELEFAX:  (317) 276-3861
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1177 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  mRNA
US-08-890-542A-3


  Query Match           16.4%;  Score 225.2;  DB 2;  Length 1177;
  Best Local Similarity   43.7%;  Pred. No. 7.5e-62;
  Matches  328;  Conservative  109;  Mismatches  303;  Indels   10;  Gaps    2;

Qy    517 TACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGAT 576
          :|||| |:| |::      |  :||||   :| |: |:|||: || || ||||||::  |:
Db    184 UACAAAAUGGAUUACCCUGAAAUGGGCUUGUGUAUAAUAAUUAAUAAUAAGAACUUCCAU 243

Qy    577 AAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAG 636
          |||   || || |:|   | :|| || || || ||    ||:|| |     |:|    ||
Db    244 AAAAGCACUGGAAUGUCAGCUCGCAAUGGUACCGAUGUCGAUGCAGCUAACCUCAGAGAG 303

Qy    637 TGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATG 696
            ::|     ||:|   :  || |: |       | ||:|||    |: |:|   | |:
Db    304 ACAUUCAUGGCCCUGAAAUACGAAGUCAGGAAUAAAAAUGACCUUACUCGUGAAGAAAUU 363

Qy    697 CAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGC 756
            ||  :| :  | |  | ::|: | || || || |  ||      | | :: |  :|
Db    364 AUGGAAUUGAUGGAUAGUGUUUCUAAGGAAGAUCACAGCAAAAGGAGCAGUUUUGUGUGU 423

Qy    757 ATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATA 816
            :  :  :||| ||:||||| |||   |:||::: :|| |   | ||           :
Db    424 GUGAUUCUAAGUCAUGGAGAUGAAGGAGUAAUUUUUGGAACGAACGGACCUGUGGACCUG 483

Qy    817 AAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAA 876
          ||  |  : ||      |:: || || ||    :||   |  |:    |  || || |||
Db    484 AAAAAACUAACUAGUUUCUUCAGAGGCGACUACUGCCGGAGUCUGACUGGAAAGCCGAAA 543

Qy    877 CTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCG 936
          |:|::| :||::||||| :|||||| || ||||: ||    || |:  || | |||
Db    544 CUCUUCAUCAUUCAGGCCUGCCGAGGUACAGAGCUGGACUGCGGUAUUGAGACAGACAGU 603

Qy    937 GGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTC 996
          ||  |   |      ||:| :|   |  |  | ||||: ||||:||  || |||::|
Db    604 GGAACUGAC---------GAUGAUAUGGCAUGCCAGAAGAUACCAGUGGGGGCCGACUUC 654

Qy    997 CTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC 1056
          |: :  || :| :| || |  || || :| :| :| :|||| |   || |    || :|
Db    655 CUGUAUGCUUACUCUACCGCACCCGGUUACUAUUCCUGGAGAAAUUCAAGGGACGGGUCA 714

Qy   1057 TGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATG 1116
          :||::  : ||  | |: :|| |||: |:| |   | |||   |  | |:|||| :||:|
Db    715 UGGUUCAUCCAGUCACUUUGCGCCAUGCUGAAACUGUACGCGCACAAGCUGGAAUUCAUG 774

Qy   1117 CAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGA 1176
          || |:||: |||  | |: ||    |  |: |||| | | :::|||:|    :|    ||
Db    775 CACAUCCU-CACUCGUGUUAACCGGAAGGUGGCCAUGGAAUUUGAGUCCUUCUCCCUGGA 833
```

```
Qy   1177 CCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACT 1236
          | |    ::||| |  ||||| |||||||| :|: : |: :| |:|||:||| || ||||:
Db    834 CGCCACUUUCCACGCAAAGAAACAGAUCCCGUGUAUUGUGUCAAUGCUCACAAAAGAACU 893

Qy   1237 CTACTTCAGTCAATAGCCATATCAGGGGTA 1266
           :||::   :|| :|     |: |  || |
Db    894 GUACUUUUAUCACUAAAGGAAUGACUGGGA 923


RESULT  14
US-08-773-608A-1
; Sequence 1, Application US/08773608A
; Patent No. 5858778
;  GENERAL INFORMATION:
;    APPLICANT:  Alnemri, Emad S.
;    APPLICANT:  Fernandes-Alnemri, Teresa
;    APPLICANT:  Litwack, Gerald
;    TITLE OF INVENTION:  Sf Caspase-1, Compositions and Methods
;    TITLE OF INVENTION:  for Making and Methods of Using the Same
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5858778ris
;      STREET:  One Liberty Place 46th floor
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  US
;      ZIP:  19355
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/773,608A
;      FILING DATE:  27-DEC-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeLuca, Mark
;      REGISTRATION NUMBER:  33,229
;      REFERENCE/DOCKET NUMBER:  TJU-2087
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2658 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  140..1039
US-08-773-608A-1


  Query Match             10.1%;  Score 139;  DB 2;  Length 2658;
  Best Local Similarity   52.6%;  Pred. No. 4.8e-34;
  Matches  399;  Conservative    0;  Mismatches  350;  Indels   10;  Gaps    4;

Qy    513 TCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTT 572
          | | ||||||||||||   | || | |  || |    ||| ||  | ||| || || | ||
Db    292 TTATTACAACATGAACCATAAACATCGTGGTATGGCCATTATTTTCAACCACGAGCATTT 351

Qy    573 TGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTT 632
            || |   |     |  ||      || |  || ||  |   |||  |||    || |
Db    352 CGACA---TTCACAGCCTGAAGTCCCGTACTGGCACGAATGTAGACAGCGACAACCTTTC 408

Qy    633 CAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAA 692
          |||     |   || | |||| ||  | || |  || |    | |||     |  |   |
Db    409 CAAAGTTCTTAAAACCTTGGGATTCAAAGTTACGGTGTTCCCTAACTTAAAATCAGAAGA 468

Qy    693 GATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGC 752
```

```
          || |  | |  |  |  |  | || | ||   ||||||| || | || | ||| ||
Db    469 AATCAATAAATTTATCCAGCAGACTGCAGAGATGGACCATTCAGACGCTGACTGTTTACT 528

Qy    753 CTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACC 812
              | | |||| ||||||||||   |    |  | ||||  ||||||  |
Db    529 TGTTGCCGTTTTAACCCATGGAGAGCTAGGAATGCTGTATGCCAAAGATACTCATTACAA 588

Qy    813 AATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACC 872
            | |  |  ||      | ||||||  |  || | ||||   || ||    | |||||
Db    589 GCCAGACAACCTGTGGTACTACTTTACTGCTGACAAATGCCCCACACTGGCTGGAAAACC 648

Qy    873 CAAACTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGA 932
          || |||||||| |||||||||||||| |||   |    | |||| ||| ||
Db    649 TAAGCTCTTCTTTATTCAGGCTTGCCAAGGTGACAGATTGGATGGTGGTATTACTCTGAG 708

Qy    933 CTCGGGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGA 992
          |          |   ||| | ||  | |||   ||   || | ||| || ||  | |||||
Db    709 CCGCACTGAGACAGATGGCTCACCTTCTA---CTTCGTATAGGATTCCCGTACATGCTGA 765

Qy    993 CTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGG 1052
          ||| |  | || | ||| || || || || || |  || |||||||  | |    |||||
Db    766 TTTCTTGATTGCATTTTCAACTGTACCTGGATACTTTTCTTGGAGGAACACTACCAGAGG 825

Qy    1053 CTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAAT 1112
           || |||||| |||||||  | ||          ||      |    |||| |||   ||| ||
Db    826 TTCATGGTTTATGCAAGCTTTGTGTGAAGAGCTACGCTATGCAGGAACAGAGAGGGACAT 885

Qy    1113 CATGCAAATCCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTG 1172
           ||  | | ||  ||     ||| | | |  ||||     ||||||||||| |  |
Db    886 TCTGACATTGCT-GACTTTTGTGTGTCAGAAGGTTGCGCTAGACTTTGAGTCTAATGCCC 944

Qy    1173 ATGACCCA---CACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCA 1229
           |||| ||       | ||| |  |||| ||  | |||||  |   |  |||||||||
Db    945 CTGACTCAGCGATGATGCATCAACAGAAACAAGTTCCCTGCATCACCAGCATGCTCACAC 1004

Qy    1230 AGGAACTCTACTTCAGTCAATAGCCATATCAGGGGTACA 1268
             ||    ||  || |  |||  |  ||  | || |
Db    1005 GTCTGCTTGTGTTTGGTAAGAAGCAGTCCCACTGATATA 1043


RESULT  15
US-08-522-813-1
; Sequence 1, Application US/08522813
; Patent No. 6033848
;  GENERAL INFORMATION:
;    APPLICANT:  Braxton, Scott M
;    APPLICANT:  Diep, Dinh
;    APPLICANT:  Seilhamer, Jeffrey J.
;    TITLE OF INVENTION:  HUMAN ICE HOMOLOG
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3330 Hillview Avenue
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/522,813
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Luther, Barbara J.
;      REGISTRATION NUMBER:  33954
;      REFERENCE/DOCKET NUMBER:  PF-0045P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-852-0195
```

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1183 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY: Lung
; CLONE: 124985
US-08-522-813-1

Query Match 7.7%; Score 105.6; DB 3; Length 1183;
Best Local Similarity 48.5%; Pred. No. 1.3e-23;
Matches 381; Conservative 0; Mismatches 374; Indels 31; Gaps 2;

```
Qy    505 CCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAAC 564
          ||  || |  |||||| ||| |       | |   || |      | ||  | ||  |
Db     52 CCGGCAGAAAAGTACAAAATGGACCACAGGAGGAGAGGAATTGCTTTAATCTTCAATCAT 111

Qy    565 AAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAG 624
          |||  ||   |        |    ||   |   |    || ||   |  ||||   ||
Db    112 GAGAGGTTCTTTTGGCACTTAACACTGCCAGAAAGGCGGGGCACCTGCGCAGATAGAGAC 171

Qy    625 GCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCT 684
             ||   |    | ||   |   || || ||||| ||||    || | ||||
Db    172 AATCTTACCCGCAGGTTTTCAGATCTAGGATTTGAAGTGAAATGCTTTTATGATCTTAAA 231

Qy    685 TGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCC 744
          |   |  | |         || |  | |  ||     |    ||| || ||||||
Db    232 GCAGAAGAACTACTGCTCAAAATTCATGAGGTGTCAACTGTTAGCCACGCAGATGCCGAT 291

Qy    745 TGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGT 804
          ||||| |  ||  || || | |||||||| ||||  |||   |||||||   | |||| |
Db    292 TGCTTTGTGTGTGTCTTCCTGAGCCATGGCGAAGGCAATCACATTTATGCATATGATGCT 351

Qy    805 GTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTA 864
             |   |||  ||   ||  | | |   || |  || || |  ||  | | |||  |
Db    352 AAAATCGAAATTCAGACATTATCTGGCTCGTTCAAAGGAGACAAGTGTCACAGCCTGGTT 411

Qy    865 GAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATG-------- 916
          |  |||||||  | ||  |||||||||| || || || | | |||  ||||
Db    412 GGAAAACCCAAGATATTTATCATTCAGGCATGTCGGGGAAACCAGCACGATGTGCCAGTC 471

Qy    917 ----------------------ATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACA 954
                                 ||   |   ||          |||  | |||
Db    472 ATTCCTTTGGATGTAGTAGATAATCAGACAGAGAAGTTGGACACCAACATAACTGAGGTG 531

Qy    955 GATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACG 1014
          |||||     |    |||| | | || || | ||||||||||||| |    || ||
Db    532 GATGCAGCCTCCGTTTACACGCTGCCTGTTGGAGCTGACTTCCTCATGTGTTACTCTGTT 591

Qy   1015 GTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTC 1074
          |    ||| ||||| ||     ||    | |  |  || || ||||   | ||||   |
Db    592 GCAGAAGGATATTATTCTCACCGGGAAACTGTGAACGGTTCATGGTACATTCAAGATTTG 651

Qy   1075 TGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGT 1134
          ||    ||  |||   |  | ||       | | ||  |||   || |||| |||   |||
Db    652 TGTGAGATGTTGGGAAAATATGGCTCCTCCTTAGAGTTCACAGAACTCCT-CACACTGGT 710

Qy   1135 GAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAA 1194
          |||    | |||| |   || |   | |         | ||||||       ||  ||
Db    711 GAACAGGAAAGTTTCTCAGCGCCGAGTGGACTTTGGCAAAGACCCAAGTGCAATTGGAAA 770

Qy   1195 GAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCC 1254
          ||||||| | |||||| ||| ||| ||||| || ||  | ||  | |||  || | |  |
Db    771 GAAGCAGGTTCCCTGTTTGGGCTCAATGCTAACTAAAAAGCTGCATTTCTTTCCAAAATC 830

Qy   1255 ATATCA 1260
             || |
Db    831 TAATTA 836
```

Search completed: August 15, 2002, 23:59:25
Job time: 6171 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2002, 21:29:14 ; Search time 2425.75 Seconds
(without alignments)
7628.290 Million cell updates/sec

Title: US-09-613-508-1
Perfect score: 1371
Sequence: 1 GCACGAGAAACTTTGCTGTG.....TAATTTCAGGAAAATGTATT 1371

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	949	69.2	973	9	AL542189	AL542189 AL542189
2	765.4	55.8	926	10	BI821936	BI821936 603039745
3	752.2	54.9	1075	10	BM459840	BM459840 AGENCOURT
4	731.4	53.3	745	10	BI824214	BI824214 603040584
5	712.2	51.9	774	9	AU130224	AU130224 AU130224
6	662.4	48.3	791	9	AU134778	AU134778 AU134778
7	651.8	47.5	719	9	AU143139	AU143139 AU143139
8	642.4	46.9	830	10	BG025549	BG025549 602274928
9	636.6	46.4	809	9	AU126220	AU126220 AU126220
10	629.8	45.9	961	10	BE891855	BE891855 601435409
11	628.2	45.8	656	10	BF971443	BF971443 602272969
12	574.4	41.9	621	10	BE894817	BE894817 601436392
13	571.4	41.7	700	9	AU117167	AU117167 AU117167
14	518.8	37.8	630	10	BE893968	BE893968 601437754
15	516	37.6	917	10	BG674973	BG674973 602621183
16	514.2	37.5	532	10	BM150092	BM150092 TCAAP4Q11
17	505	36.8	796	10	BE619147	BE619147 601472883
18	503	36.7	1006	10	BG339524	BG339524 602437368
19	482	35.2	497	10	BM194512	BM194512 TCAAP2Q11
20	478.8	34.9	923	10	BF698196	BF698196 602125529
21	471.4	34.4	485	10	BM147460	BM147460 TCAAP1Q11
22	465	33.9	777	9	AI115889	AI115889 ue96b06.y
23	461.6	33.7	984	10	BI692031	BI692031 603311680
24	443.4	32.3	774	10	BG966729	BG966729 602834592
25	441.8	32.2	787	9	AI315897	AI315897 uj27d07.y
26	432.4	31.5	697	10	BF786558	BF786558 602108505
27	420	30.6	846	10	BF668893	BF668893 602123296
28	418	30.5	808	10	BF540869	BF540869 602067587
29	410	29.9	756	10	BI458770	BI458770 603200030
30	402.6	29.4	870	10	BG404254	BG404254 602420322
31	400.6	29.2	840	9	AU131461	AU131461 AU131461
32	375.6	27.4	588	10	BI401667	BI401667 MI-P-CP0-
33	362.8	26.5	402	9	AA101428	AA101428 zn72e08.r
c 34	359.6	26.2	486	9	AI909733	AI909733 IL-BT223-
c 35	332.2	24.2	437	10	BE749735	BE749735 200771 MA
36	313	22.8	558	9	AI119367	AI119367 ue98g06.y
37	308.8	22.5	479	10	BI080773	BI080773 602878538
38	292.6	21.3	412	9	AA939788	AA939788 vz92g09.r
39	290.4	21.2	473	9	BB862907	BB862907 BB862907
40	289	21.1	290	9	AA297423	AA297423 EST112974
41	287.2	20.9	394	10	BG918158	BG918158 602821076
42	278.4	20.3	356	9	AA298267	AA298267 EST114122
43	277.2	20.2	553	10	BE380736	BE380736 601270846
44	269	19.6	593	10	H91868	H91868 ys81a06.r1
45	261.4	19.1	360	10	BG967610	BG967610 602833403

ALIGNMENTS

RESULT 1
AL542189
LOCUS AL542189 973 bp mRNA linear EST 16-FEB-2001
DEFINITION AL542189 LTI_FL002_PL1 Homo sapiens cDNA clone CS0DE008YG21 5 prime , mRNA sequence.
ACCESSION AL542189
VERSION AL542189.1 GI:12873988
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 973)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES Location/Qualifiers
source 1. .973
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DE008YG21"
/clone_lib="LTI_FL002_PL1"
/lab_host="DH10B"
/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville , Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 275 a 239 c 241 g 217 t 1 others
ORIGIN

```
  Query Match            69.2%;  Score 949;  DB 9;  Length 973;
  Best Local Similarity  99.6%;  Pred. No. 1.3e-256;
  Matches  971;  Conservative    1;  Mismatches    1;  Indels    2;  Gaps    2;

Qy    305 TGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGGGCTGTATTG 364
          ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      1 TGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGA-TCAGGGCTGTATTG 59

Qy    365 AAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGT 424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 AAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGT 119

Qy    425 CCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCA 484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 CCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCA 179

Qy    485 AGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCA 544
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 AGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCA 239

Qy    545 AATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACG 604
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 AATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACG 299

Qy    605 GAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGA 664
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 GAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGA 359

Qy    665 TTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAG 724
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 TTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAG 419

Qy    725 AGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATG 784
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 AGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATG 479

Qy    785 TAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGG 844
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 TAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGG 539

Qy    845 ATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGA 904
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 ATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGA 599

Qy    905 CCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATC 964
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 CCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATC 659

Qy    965 CTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCT 1024
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 CTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCT 719

Qy   1025 ATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCC 1084
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 ATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCC 779

Qy   1085 TGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGACAGA 1144
          |||||||||||||||||||||||||||||| |||| ||||||||||||||||||||||||
Db    780 TGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCAGGGTGAATGACAGA 838

Qy   1145 GTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGATC 1204
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    839 GTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGATC 898

Qy   1205 CCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCAGGGG 1264
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    899 CCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCAGGGG 958

Qy   1265 TACATTCTAGCTGAG 1279
          :|||||||||||||||
Db    959 KACATTCTAGCTGAG 973
```

```
RESULT   2
BI821936
LOCUS       BI821936                926 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603039745F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5180651 5',
            mRNA sequence.
ACCESSION   BI821936
VERSION     BI821936.1  GI:15933486
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 926)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11450  row: j  column: 12
            High quality sequence stop: 882.
FEATURES             Location/Qualifiers
     source          1. .926
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5180651"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69.  Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon  cloning).  Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was  constructed by C.
                     Gruber (Invitrogen).  Research Genetics tracking  code
                     021. Note: this is a NIH_MGC Library."
BASE COUNT      264 a    218 c    242 g    202 t
ORIGIN
```

```
  Query Match            55.8%;  Score 765.4;  DB 10;  Length 926;
  Best Local Similarity  95.6%;  Pred. No. 8.9e-205;
  Matches  884;  Conservative    0;  Mismatches   31;  Indels   10;  Gaps    9;

Qy    287 AGCGACGGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGA 346
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AGCGACGGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGA 60

Qy    347 TGATTCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGG 406
          |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGA-GCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGG 119

Qy    407 ATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATG 466
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 ATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATG 179

Qy    467 TCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTAC-ATATCAGTACAACATG 525
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    180 TCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACGATATCAGTACAACATG 239

Qy    526 AATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACA 585
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 AATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACA 299
```

```
Qy     586 GGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGA 645
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 GGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGA 359

Qy     646 AGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTG 705
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 AGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTG 419

Qy     706 CTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTA 765
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     420 CTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTA 479

Qy     766 AGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTG 825
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     480 AGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTG 539

Qy     826 ACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTC 885
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     540 ACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTC 599

Qy     886 ATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATC 945
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 ATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATC 659

Qy     946 AATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCC 1005
           ||||||||||||||||||||||||||||||||    | |     ||||||||||||
Db     660 AATGACACAGATGCTAATCCTCGATACAAGATCCAGTGGAAGCTGACTTTCCTCTTCGCC 719

Qy    1006 TATTCCACGGTTCCA-GGCTATTACTCGTGGAGGAGCCCAGGAAGA-GGCTCCTGGTTTG 1063
           ||||||||||||||| |||||||||||||||||||||||||||||| |||||||||||| |
Db     720 TATTCCACGGTTCCAGGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGGCTCCTGGTTCG 779

Qy    1064 TGCAAGCCCTCT--GCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAAT 1121
           ||||||||||||     |  ||||||||||||||||||||| |||||||||||||  ||
Db     780 TGCAAGCCCTCTTGCTCCATTCCTGGAGGAGCACGGAAAAGA-CTGGAAATCATGCAGAT 838

Qy    1122 CCTCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCAC 1181
           ||| ||| ||||||||||||||| || ||||||||||||||||||||||||  || |
Db     839 CCT-CACAAGGGTGAATGACAGAG-TGGCAGGCACTTTGAGTCTCAGTCTGATGGACC-C 895

Qy    1182 ACTTCCATGAGAAGAAGCAGATCCC 1206
           | ||||    |||| | ||||||||
Db     896 AATTCCTGAGGAAGGAACAGATCCC 920
```

RESULT 3
BM459840
LOCUS BM459840 1075 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6422147 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5532181 5', mRNA sequence.
ACCESSION BM459840
VERSION BM459840.1 GI:18508880
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1075)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12215 row: a column: 14
High quality sequence stop: 545.
FEATURES Location/Qualifiers
source 1. .1075
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5532181"
/clone_lib="NIH_MGC_71"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb. "
BASE COUNT 296 a 271 c 271 g 234 t 3 others
ORIGIN

Query Match 54.9%; Score 752.2; DB 10; Length 1075;
Best Local Similarity 89.2%; Pred. No. 5.1e-201;
Matches 867; Conservative 0; Mismatches 71; Indels 34; Gaps 4;

```
Qy     289 CGACGGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATG 348
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CGACGGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATG 60

Qy     349 ATTCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGAT 408
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 A-GCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGAT 119

Qy     409 GCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTT---------------------- 446
           ||||||||||||||||||||||||||||||||||||
Db     120 GCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGCCCCTGACTCTGGAACTTT 179

Qy     447 ---------CAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGA 497
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180 ATATTTCACCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGA 239

Qy     498 CCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAAT 557
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     240 CCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAAT 299

Qy     558 AAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGA 617
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 AAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGA 359

Qy     618 TGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGA 677
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 TGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGA 419

Qy     678 CTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAA 737
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     420 CTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAA 479

Qy     738 TGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAA 797
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     480 TGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAA 539

Qy     798 AGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAAC 857
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     540 AGATGGTGTCACACCAATANAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAAC 599

Qy     858 CCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGA 917
           |||||||||||||||||||||  |||||||||||||||||| | ||||||||||||
Db     600 CCTTTTAGAGAAACCCAAACTCTTTCTCATTCAGGCTTGCCGAGGNAACGAGCTTGATGA 659

Qy     918 TGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGAT 977
           |||||||||| |||||||||| |||||||||||||||||||||||||||||||| ||||
Db     660 TGGCATCCAGGNCGACTCGGGGGCCATCAATGACACAGATGCTAATCCTCGATAC-AGAT 718

Qy     978 CCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAG 1037
           ||||||||| |||||||||||| |||| |  |||| |||||| ||||||||||
Db     719 CCCAGTGGAAACTGACTTCCTCTTCCCCTATCCACGGGTTCAAGGCTATAACTCGTGGAG 778

Qy    1038 GAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGG 1097
           || |||||  ||| |||||||||| |  |||||||||||||| ||  |  | | ||||
Db     779 GAAGCCAGGAGAAGGTTCCTGGTTTGGGGCAGCCCTCTGCTCCTTCTGGAGGAAACACGG 838
```

```
Qy   1098 AAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGACAGAGTT-GCCAGGCAC 1156
          ||||||||||||||| || |||||||||||| |  ||    |   ||  |||| ||||||
Db    839 AAAAGACCTGGAATTCTTGCAAATCCTCCCCAGGGTGAATGAACGAAGTTGGCCAGGACT 898

Qy   1157 TTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGATCCCCTGTGTGGTC 1216
          ||||||  |||| | || ||||| | ||| | ||| |||||||  | |||| | | ||
Db    899 TTTGAGGTTCAGCCGGAGGACCCCCCCTTTCCTGAAAAGAAGCCAAACCCCGGGGGGGGG 958

Qy   1217 TCCATGCTCACC 1228
          ||  |  | ||
Db    959 CCCCCGGGCCCC 970


RESULT   4
BI824214
LOCUS       BI824214                 745 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603040584F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5181381 5',
            mRNA sequence.
ACCESSION   BI824214
VERSION     BI824214.1  GI:15935764
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 745)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11452  row: h  column: 22
            High quality sequence stop: 740.
FEATURES             Location/Qualifiers
     source          1. .745
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5181381"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69.  Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon  cloning).  Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was  constructed by C.
                     Gruber (Invitrogen).  Research Genetics tracking  code
                     021. Note: this is a NIH_MGC Library."
BASE COUNT      215 a    178 c    185 g    167 t
ORIGIN


  Query Match           53.3%;  Score 731.4;  DB 10;  Length 745;
  Best Local Similarity   99.7%;  Pred. No. 3.2e-195;
  Matches  743;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy    298 GAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGGGC 357
          ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db      1 GAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGA-TCAGGGC 59

Qy    358 TGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCA 417
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 TGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCCA 119

Qy    418 GACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGA 477
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 GACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGA 179

Qy    478 TCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAG 537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 TCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAG 239

Qy    538 CTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTT 597
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 CTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGTT 299

Qy    598 CGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTT 657
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 CGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTT 359

Qy    658 GACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCT 717
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 GACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCT 419

Qy    718 TCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAA 777
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 TCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAA 479

Qy    778 GAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTT 837
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 GAACATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTT 539

Qy    838 AGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGC 897
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 AGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGC 599

Qy    898 CGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGAT 957
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 CGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGAT 659

Qy    958 GCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTT 1017
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 GCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTT 719

Qy   1018 CCAGGCTATTACTCGTGGAGGAGCC 1042
          |||||||||||||||||||||||||
Db    720 CCAGGCTATTACTCGTGGAGGAGCC 744


RESULT   5
AU130224
LOCUS       AU130224                 774 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  AU130224 NT2RP3 Homo sapiens cDNA clone NT2RP3000452 5', mRNA
            sequence.
ACCESSION   AU130224
VERSION     AU130224.1  GI:10990578
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 774)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
```

Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

```
FEATURES             Location/Qualifiers
     source          1. .774
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP3000452"
                     /clone_lib="NT2RP3"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                     cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT      220 a    179 c    193 g    178 t      4 others
ORIGIN
```

Query Match 51.9%; Score 712.2; DB 9; Length 774;
Best Local Similarity 98.8%; Pred. No. 8.6e-190;
Matches 758; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

```
Qy    237 GAGACTTTTAGTTTCGCTTTCGCTAAAGGGGCCCCAGACCCTTGCTGCGGAGCGACGGAG 296
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAGACTTTTAGTTTCGCTTTCGCTAAAGGGGCCCCAGACCCTTGCTGCGGAGCGACGGAG 60

Qy    297 AGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGATTCAGGG 356
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     61 AGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATGGCAGATGA-TCAGGG 119

Qy    357 CTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCC 416
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 CTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTCAGTGGATGCTAAGCC 179

Qy    417 AGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCG 476
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 AGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCG 239

Qy    477 ATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAA 536
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 ATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAA 299

Qy    537 GCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGT 596
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 GCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGTATGGGCGT 359

Qy    597 TCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTT 656
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 TCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTT 419

Qy    657 TGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGC 716
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 TGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGC 479

Qy    717 TTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGA 776
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 TTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGA 539

Qy    777 AGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTT 836
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 AGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTT 599

Qy    837 TAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTG 896
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db    600 TAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCANGCTTG 659

Qy    897 CCGAGGGACCGAGCTTGATGATGGCATCCAGG-CCGACTCGGGGCCCATCAATGACACAG 955
          |||||||||||||||||||||||||| |||| ||||||||||||||||||||| |||
Db    660 CCGAGGGACCGAGCTTGATGATGGCATNCAGGCCCGACTCGGGGCCCATCAATGAC-CAG 718

Qy    956 ATGCTAATCCTCGATACAAGATCCCAGTGG-AAGCTGACTTCCTCTT 1001
          |||||||||| |||||||||||||||||| |||||||||| || ||
Db    719 ATGCTAATCCTTGATACAAGATCCCAGTGGNAAGCTGACTTTCTTTT 765
```

```
RESULT   6
AU134778
LOCUS       AU134778                791 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  AU134778 PLACE1 Homo sapiens cDNA clone PLACE1000485 5', mRNA
            sequence.
ACCESSION   AU134778
VERSION     AU134778.1  GI:10995317
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 791)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .791
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="PLACE1000485"
                     /clone_lib="PLACE1"
                     /tissue_type="placenta"
                     /note="Vector: pME18SFL3"
BASE COUNT      234 a    189 c    195 g    166 t      7 others
ORIGIN
```

Query Match 48.3%; Score 662.4; DB 9; Length 791;
Best Local Similarity 98.4%; Pred. No. 1e-175;
Matches 688; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

```
Qy    272 AGACCCTTGCTGCGGAGCGACGGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGT 331
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      1 AGACCCTTGCTGCGGAGCGACNGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGT 60

Qy    332 GGGAACGATGGCAGATGATTCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAA 391
          |||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     61 GGGAACGACGGCAGATGA-TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAA 119

Qy    392 ATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTA 451
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 ATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTA 179

Qy    452 AGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACAT 511
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 AGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACAT 239

Qy    512 ATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACT 571
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 ATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACT 299

Qy    572 TTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCT 631
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    300 TTGATAAAGTGACAGGTTTGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCT 359

Qy    632 TCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCA 691
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 TCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCA 419
```

```
Qy     692 AGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCG 751
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     420 AGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCG 479

Qy     752 CCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACAC 811
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     480 CCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACAC 539

Qy     812 CAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAG-AGAAA 870
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| |
Db     540 CAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAAGAAA 599

Qy     871 CCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCC 930
           ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||| |||
Db     600 CCCAAACTCTTCTTCATTCANGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAAGCC 659

Qy     931 GACTCGGGGCCCATCAATGACACAGATGCTAATCCTCGA 969
           ||||| ||| ||||||||||||||||||||||||| | |
Db     660 GACTCCGGGGCCATCAATGACACAGATGCTAATCCCCNA 698
```

RESULT 7
AU143139
LOCUS AU143139 719 bp mRNA linear EST 25-OCT-2000
DEFINITION AU143139 Y79AA1 Homo sapiens cDNA clone Y79AA1001448 5', mRNA sequence.
ACCESSION AU143139
VERSION AU143139.1 GI:11004660
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 719)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
FEATURES Location/Qualifiers
source 1. .719
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Y79AA1001448"
/clone_lib="Y79AA1"
/cell_type="retinoblastoma"
/cell_line="Y79"
/note="Vector: pME18SFL3"
BASE COUNT 212 a 169 c 181 g 154 t 3 others
ORIGIN

Query Match 47.5%; Score 651.8; DB 9; Length 719;
Best Local Similarity 98.6%; Pred. No. 9.4e-173;
Matches 708; Conservative 0; Mismatches 5; Indels 5; Gaps 5;

```
Qy     272 AGACCCTTGCTGCGGAGCGACGGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGT 331
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 AGACCCTTGCTGCGGAGCGACGGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGT 62

Qy     332 GGGAACGATGGCAGATGATTCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAA 391
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db      63 GGGAACGATGGCAGATGA-TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAA 121

Qy     392 ATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTA 451
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 ATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTA 181

Qy     452 AGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACAT 511
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 AGAAGAAGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACAT 241

Qy     512 ATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACT 571
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     242 ATCAGTACAACATGAATTTTGAAAAGCCGGGCAAATGCATCATAATAAACAACAAGAACT 301

Qy     572 TTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCT 631
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 TTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCT 361

Qy     632 TCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCA 691
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 TCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCA 421

Qy     692 AGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCG 751
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     422 AGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCG 481

Qy     752 CCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACAC 811
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     482 CCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACAC 541

Qy     812 CAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAAC 871
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     542 CAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAAC 601

Qy     872 CCAAACTCTTCTTCATTCAGGCTTGCCGAGGGA-CCGAGCTTGATGATGGCATCCAGG-C 929
           |||||||||||||||||| |||||||||||| |||||||||||||||||||| || | |
Db     602 CCAAACTCTTCTTCATTCANGCTTGCCGAGGGACCCGAGCTTGATGATGGCATNCANGCC 661

Qy     930 CGACTCGGGGCCCATCAAT-GACACAGATGCTAATCCTCGA-TACAAGATCCCAGTGG 985
           |||||||||||||||||| |||||||||||||||||||| | ||||||||||||||
Db     662 CGACTCGGGGCCCATCAATGGACACAGATGCTAATCCTCGATTCCAAGATCCCAGTGG 719
```

RESULT 8
BG025549
LOCUS BG025549 830 bp mRNA linear EST 24-JAN-2001
DEFINITION 602274928F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4363038 5', mRNA sequence.
ACCESSION BG025549
VERSION BG025549.1 GI:12412258
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 830)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis Staudt, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10008 row: g column: 07
High quality sequence stop: 669.
FEATURES Location/Qualifiers
source 1. .830
/organism="Homo sapiens"

```
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4363038"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      229 a    206 c    196 g    197 t      2 others
ORIGIN

  Query Match             46.9%;  Score 642.4;  DB 10;  Length 830;
  Best Local Similarity   93.0%;  Pred. No. 4.6e-170;
  Matches  729;  Conservative    0;  Mismatches   43;  Indels   12;  Gaps    5;

Qy    502 GTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAAC 561
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTGCCTACATATCAGTACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAAC 60

Qy    562 AACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCC 621
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AACAAGAACTTTGATAAAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCC 120

Qy    622 GAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGC 681
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GAGGCGCTCTTCAAGTGCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGC 180

Qy    682 TCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCC 741
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCTTGTGCCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCC 240

Qy    742 GCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGAT 801
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGAT 300

Qy    802 GGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTT 861
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTT 360

Qy    862 TTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGC 921
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGC 420

Qy    922 ATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCA 981
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCA 480

Qy    982 GTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGC 1041
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGC 540

Qy   1042 CCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGG--AGCACGGAA 1099
          |||||||||||||||||||||||||||||||||||||||||||||||||      | ||
Db    541 CCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGAGGGAGCACGGCAA 600

Qy   1100 AAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGACA-GAGTTGCCAGGCACTT 1158
          |||||||||||||||||| ||||||  |  || ||||||||| ||||||||||||||||
Db    601 AAGACCTGGAAATCATGCAGATCCTCACCAGGGNTGAATGACAGGAGTTGCCAGGCACTT 660

Qy   1159 TGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAA-----GCAGATCCCCTGTGTG 1213
          ||||||||| ||||     | |||        | ||||      ||||||||||||||||
Db    661 TGAGTCTCAGGCTGAATGAACNCACACTTCCCATGAGAACGAAGCCAGATCCCCTGTGTG 720

Qy   1214 GTCTCCATGCTC---ACCAAGGAACTCTAC-TTCAGTCAATAGCCATATCAGGGGTACAT 1269
          |||| ||    |   ||| |||||||| | ||||| | |  ||||||||||||||||| |
Db    721 GTCTTCAATGGCTCACCCAGGGAACTCTTCTTTCAGCCCACTGCCATATCAGGGGTACTT 780

Qy   1270 TCTA 1273
            |||
Db    781 CCTA 784
```

```
RESULT   9
AU126220
LOCUS       AU126220                809 bp    mRNA    linear   EST 23-OCT-2000
DEFINITION  AU126220 NT2RP1 Homo sapiens cDNA clone NT2RP1000352 5', mRNA
            sequence.
ACCESSION   AU126220
VERSION     AU126220.1  GI:10950936
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 809)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Kawai,Y., Ishii,S., Saito,K.,
            Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
  TITLE     HRI human cDNA project (Ota,T., Nishikawa,T., Suzuki,Y., Kawai,Y.,
            Ishii,S., Saito,K., Nakamura,Y., Nagai,T., Sugano,S., Isogai,T.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .809
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2RP1000352"
                     /clone_lib="NT2RP1"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="Vector: pUC19FL3; mRNA from NT2 neuronal precursor
                     cells after 48-hours retinoic acid (RA) induction"
BASE COUNT      230 a    197 c    197 g    178 t      7 others
ORIGIN

  Query Match             46.4%;  Score 636.6;  DB 9;  Length 809;
  Best Local Similarity   96.7%;  Pred. No. 2e-168;
  Matches  667;  Conservative    0;  Mismatches   21;  Indels    2;  Gaps    2;

Qy    338 GATGGCAGATGATTCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAG 397
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    118 GATGGCAGATGA-TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAG 176

Qy    398 ATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGA 457
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    177 ATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGA 236

Qy    458 AGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGT 517
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    237 AGAAAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGT 296

Qy    518 ACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATA 577
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    297 ACAACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATA 356

Qy    578 AAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGT 637
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    357 AAGTGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGT 416

Qy    638 GCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGC 697
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    417 GCTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGC 476
```

```
Qy    698 AAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCA 757
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    477 AAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCA 536

Qy    758 TCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAA 817
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    537 TCCTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAA 596

Qy    818 AGGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAAC 877
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    597 ANGATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAAC 656

Qy    878 TCTTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGG 937
          |||||||||||||||||||||| ||| ||| ||||||||||||||||| |||||||| |
Db    657 TCTTCTTCATTCAGGCTTGCCGAAGGAACGANCTTGATGATGGCATCCAAGCCGACTCCG 716

Qy    938 GGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCC 997
          || |||||||||||| || |||||||||||||||||| | |||||  ||||| |||| ||||
Db    717 GGGCCATCAATGAC-CANATGCTAATCCTCGATACCANATCCCCTTGGAAACTGANTTCC 775

Qy    998 TCTTCGCCTATTCCACGGTTCCAGGCTATT 1027
          |||| |||||||| |||||||  | | ||
Db    776 NCTTCCCCTATTCCCCGGTTCCCNGGTTTT 805


RESULT  10
BE891855
LOCUS       BE891855                961 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601435409F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920522 5',
            mRNA sequence.
ACCESSION   BE891855
VERSION     BE891855.1  GI:10351599
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 961)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9752  row: a  column: 03
            High quality sequence start: 5
            High quality sequence stop: 662.
FEATURES             Location/Qualifiers
     source          1. .961
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3920522"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      274 a    224 c    254 g    209 t
ORIGIN


  Query Match           45.9%;  Score 629.8;  DB 10;  Length 961;
  Best Local Similarity   94.0%;  Pred. No. 1.8e-166;
  Matches  721;  Conservative    0;  Mismatches   37;  Indels    9;  Gaps    6;
```

```
Qy    282 TGCGGAGCGACGGA-GAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGAT 340
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGCGGAGCGACGGAGGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGAT 60

Qy    341 GGCAGATGATTCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATT 400
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGCAGATGA-TCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATT 119

Qy    401 CAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGA 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 CAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGA 179

Qy    461 AAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACA 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 AAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACA 239

Qy    521 ACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAG 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 ACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAG 299

Qy    581 TGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCT 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 TGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCT 359

Qy    641 TCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAG 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 TCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAG 419

Qy    701 ATCTGCTT-AAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATC 759
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 ATCTGCTTAAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATC 479

Qy    760 CTCTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAG 819
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    480 CTCTTAAGCCATGGAGAAGAAAATGT-ATTTATGGGAAAGATGGTGTCACACCAATAAAG 538

Qy    820 GATTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTC 879
          |||||||||||||||||||||||||  ||||||||| ||||||||||||||||||||| ||
Db    539 GATTTGACAGCCCACTTTAGGGGGGA-CGATGCAAAAACCTTTTAGAGAAACCCAAAATC 597

Qy    880 TTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGG 939
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    598 TTCTTCATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGG 657

Qy    940 CCCATCAATGACACAGA----TGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTT 995
          ||    |   ||||||       |  |||   | | |    ||  |||||| ||  ||
Db    658 CCATCAATGAGCACAGAATGCCTATCCTCCGATACAAAGATCCCAGGTGGAAGGTGGATT 717

Qy    996 CCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCC 1042
          |||||||| |||||||||| | |  | ||||||||||| ||||||
Db    718 CCTCTTCGGCTATTCCACGGGTTCCAGGTATTACTCGTGGGGGAGCC 764


RESULT  11
BF971443
LOCUS       BF971443                656 bp    mRNA    linear   EST 22-JAN-2001
DEFINITION  602272969F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4361176 5',
            mRNA sequence.
ACCESSION   BF971443
VERSION     BF971443.1  GI:12338658
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 656)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
```

```
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM10003  row: i  column: 17
           High quality sequence stop: 649.
FEATURES             Location/Qualifiers
     source          1. .656
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4361176"
                     /clone_lib="NIH_MGC_84"
                     /tissue_type="adrenal cortex carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site_1:
                     NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT
                     primed. Average insert size 1.229 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      188 a    155 c    161 g    152 t
ORIGIN


  Query Match           45.8%;  Score 628.2;  DB 10;  Length 656;
  Best Local Similarity   99.2%;  Pred. No. 4.2e-166;
  Matches  652;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    2;

Qy     525 GAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGAC 584
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGAC 60

Qy     585 AGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCG 644
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTTCCG 120

Qy     645 AAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCT 704
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGATCT 180

Qy     705 GCTTAAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTT 764
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GCTTAAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCTCTT 240

Qy     765 AAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTT 824
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTT 300

Qy     825 GACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTT 884
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTT 360

Qy     885 CATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCAT 944
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CATTCAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCAT 420

Qy     945 CAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGC 1004
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGC 480

Qy    1005 CTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCC-TGGTTTG 1063
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||| 
Db     481 CTATTCCACGGTTCCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTTGGTTTG 540

Qy    1064 TGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCC 1123
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||    
Db     541 TGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCC 600

Qy    1124 TCCACCAGGGTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCA 1180
           | ||||||| ||||||||||||||||||||||||||||||||||||||||||| ||| 
Db     601 T-CACCAGGCTGAATGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGAACCA 656
```

```
RESULT  12
BE894817
LOCUS       BE894817                621 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601436392F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3921455 5',
            mRNA sequence.
ACCESSION   BE894817
VERSION     BE894817.1  GI:10357589
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          . Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 621)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM9754  row: g  column: 24
           High quality sequence stop: 618.
FEATURES             Location/Qualifiers
     source          1. .621
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3921455"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      188 a    137 c    161 g    135 t
ORIGIN


  Query Match           41.9%;  Score 574.4;  DB 10;  Length 621;
  Best Local Similarity   99.2%;  Pred. No. 6.3e-151;
  Matches  619;  Conservative    0;  Mismatches    1;  Indels    4;  Gaps    4;

Qy     282 TGCGGAGCGACGGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATG 341
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TGCGGAGCGACGGAGAGAGACTGTGCCAGTCCCAGCCGCCCTACCGCCGTGGGAACGATG 60

Qy     342 GCAGATGATTCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTC 401
           ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GCAGATGA-GCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAGCAAATGAAGATTC 119

Qy     402 AGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAA 461
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     120 AGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCGTCCCTCTTCAGTAAGAAGAAGAA 179

Qy     462 AAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAA 521
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180 AAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACAA 239

Qy     522 CATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGT 581
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     240 CATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGT 299

Qy     582 GACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTT 641
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 GACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCTT 359

Qy     642 CCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGA 701
           |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
```

```
Db     360 CCGAAGCCTGGGTTTTGACGTGA-TGTCTATAATGACTGCTCTTGTGCCAAGATGCAAGA 418

Qy     702 TCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCT 761
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     419 TCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCCT 478

Qy     762 CTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGA 821
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     479 CTTAAGCCATGGAGAAGAAAATGTAATTTATGGGAAAGATGGTGTCACACCAATAAAGGA 538

Qy     822 TTTGACAGCCCACTTTAGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTT 881
           |||||||||||||||||||||||||||||| |||||||||||||||||| ||||
Db     539 TTTGACAGCCCACTTTAGGGGGGATAGATGCAAAA-CCTTTTAGAGAAACCCAAA-TCTT 596

Qy     882 CTTCATTCAGGCTTGCCGAGGGAC 905
           |||||||||||||||||||||||
Db     597 CTTCATTCAGGCTTGCCGAGGGAC 620
```

```
RESULT  13
AU117167
LOCUS       AU117167                 700 bp    mRNA    linear   EST 19-OCT-2000
DEFINITION  AU117167 HEMBA1 Homo sapiens cDNA clone HEMBA1000808 5', mRNA
            sequence.
ACCESSION   AU117167
VERSION     AU117167.1  GI:10932113
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 700)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .700
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEMBA1000808"
                     /clone_lib="HEMBA1"
                     /tissue_type="whole embryo, mainly head"
                     /dev_stage="embryo, 10 weeks"
                     /note="Vector: pME18SFL3"
BASE COUNT      188 a    158 c    154 g    197 t      3 others
ORIGIN
```

```
  Query Match           41.7%;  Score 571.4;  DB 9;  Length 700;
  Best Local Similarity  99.7%;  Pred. No. 4.7e-150;
  Matches  583;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy     787 ATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGAT 846
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATTTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTAGGGGGGAT 60

Qy     847 AGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACC 906
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGCCGAGGGACC 120
```

```
Qy     907 GAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCT 966
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGATGCTAATCCT 180

Qy     967 CGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTAT 1026
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTAT 240

Qy    1027 TACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTG 1086
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTG 300

Qy    1087 GAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAATGACAGAGT 1146
           |||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||||
Db     301 GAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCAGGGTGAATGACAGAGT 359

Qy    1147 TGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGATCCC 1206
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 TGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAAGCAGATCCC 419

Qy    1207 CTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCAGGGGTA 1266
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     420 CTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATATCAGGGGTA 479

Qy    1267 CATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGCTCTTGA 1326
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     480 CATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTATGCTCTTGA 539

Qy    1327 AATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
           |||||||||||||||||||||||||||||||||||||||||||||
Db     540 AATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 584
```

```
RESULT  14
BE893968
LOCUS       BE893968                 630 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601437754F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3922716 5',
            mRNA sequence.
ACCESSION   BE893968
VERSION     BE893968.1  GI:10355868
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 630)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9757  row: 1  column: 13
            High quality sequence start: 4
            High quality sequence stop: 566.
FEATURES             Location/Qualifiers
     source          1. .630
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3922716"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
```

```
BASE COUNT      166 a    156 c    141 g    167 t
ORIGIN

  Query Match             37.8%;  Score 518.8;  DB 10;  Length 630;
  Best Local Similarity   99.4%;  Pred. No. 3.2e-135;
  Matches  531;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy     838 AGGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGC 897
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ACGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATTCAGGCTTGC 60

Qy     898 CGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGAT 957
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGGGCCCATCAATGACACAGAT 120

Qy     958 GCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTT 1017
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCTAATCCTCGATACAAGATCCCAGTGGAAGCTGACTTCCTCTTCGCCTATTCCACGGTT 180

Qy    1018 CCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGC 1077
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CCAGGCTATTACTCGTGGAGGAGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGC 240

Qy    1078 TCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAA 1137
           |||||||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||
Db     241 TCCATCCTGGAGGAGCACGGAAAAGACCTGGAAATCATGCAGATCCT-CACCAGGGTGAA 299

Qy    1138 TGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAA 1197
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 TGACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCATGAGAAGAA 359

Qy    1198 GCAGATCCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATA 1257
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 GCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCTACTTCAGTCAATAGCCATA 419

Qy    1258 TCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTA 1317
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     420 TCAGGGGTACATTCTAGCTGAGAAGCAATGGGTCACTCATTAATGAATCACATTTTTTTA 479

Qy    1318 TGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 1371
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     480 TGCTCTTGAAATATTCAGAAATTCTCCAGGATTTTAATTTCAGGAAAATGTATT 533


RESULT  15
BG674973
LOCUS       BG674973                 917 bp    mRNA    linear   EST 01-MAY-2001
DEFINITION  602621183F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4746621 5',
            mRNA sequence.
ACCESSION   BG674973
VERSION     BG674973.1  GI:13906369
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 917)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
            Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10595  row: e  column: 22
            High quality sequence stop: 715.
FEATURES             Location/Qualifiers
     source          1. .917
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4746621"
                     /clone_lib="NCI_CGAP_Skn3"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5kb.  Library constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      263 a    251 c    202 g    201 t
ORIGIN

  Query Match             37.6%;  Score 516;  DB 10;  Length 917;
  Best Local Similarity   93.9%;  Pred. No. 2.4e-134;
  Matches  712;  Conservative    0;  Mismatches   25;  Indels   21;  Gaps   16;

Qy     461 AAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACA 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 AAAATGTCACCATGCGATCCATCAAGACCACCCGGGACCGAGTGCCTACATATCAGTACA 60

Qy     521 ACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAG 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACATGAATTTTGAAAAGCTGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAG 120

Qy     581 TGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCT 640
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGACAGGTATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTGCT 180

Qy     641 TCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAG 700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTGCCAAGATGCAAG 240

Qy     701 ATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCC 760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAATGCCGCCTGCTTCGCCTGCATCC 300

Qy     761 TCTTAAGCCATGG-AGAAGAAAATGTAATTTATGGGAAAGA-TGGTGTCACACCAA--TA 816
           |||||||||||| |||||||||||||||||||||||||||| |||||||||||||   |
Db     301 TCTTAAGCCATGGCAGAAGAAAATGTAATTTATGGGAAAGACTGGTGTCACACCAACTAA 360

Qy     817 AAGGATTTGACAGCCCACTTTAGGGGGG---ATAGATGCAAAACCCTTT-TAGAGAAACC 872
           | |  |||||||||||||| |  ||| |   |||||||||||||||||| |||||||||
Db     361 AGGACTTTGACAGCCCACTCTTAGGGCGGCGATAGATGCAAAACCCTTTCTAGAGAAACC 420

Qy     873 CAAACTCTTCTTCATTCAGGCT-TGCCGAGGGACCGAGCTTGATG-ATGGCATCCAGGCC 930
           ||||||||||||||||||||| ||||||||||||||||||||||| ||||||||||||||
Db     421 CAAACTCTTCTTCATTCAGGCTATGCCGAGGGACCGAGCTTGATGCATGGCATCCAGGCC 480

Qy     931 GACTCGGGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCT 990
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 GACTCGGGGCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAGCT 540

Qy     991 GACTTCCTCTTCGCCTATTCCACGGT-TCCAGGCTATTACTCGTG-GAGGAGCCCAGGAA 1048
           ||||||||||||||||||||||||| ||||||||||||||||||| ||||||||||||||
Db     541 GACTTCCTCTTCGCCTATTCCACGGTCTCCAGGCTATTACTCGTGCGAGGAGCCCAGGAA 600

Qy    1049 -GAGGCTCCTGGTTT--GTGCAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAA--AAGA 1103
            |||||||||||||       |||||||||||||||||||||||||||||||  |
Db     601 CGAGGCTCCTGGTTTTGTGCAAAGCCCTCTGCTCCATCCTGGAGGAGCACGGAAACAGAG 660

Qy    1104 CCTGGAAATCATGCAAATCCTCCACCAGGGTGAA-TGACAGAG-TTGCCAGGCACTTTG- 1160
           |  | || |||||||||||||  |  |||||| |||| || ||||||||||||||||
Db     661 CTGGAAACTCATGCAAATCCTCACCAGCGGTGAACTGACACAGTTTGCCAGGCACTTTGA 720

Qy    1161 AGTCTCAGTCTG-ATGACCCACACTTCCATGAGAAGAA 1197
           |||||||||||| |||||||||||||||||||||||||
Db     721 AGTCTCAGTCTGAATGACCCACACTTCCATGAGAAGAA 758


Search completed: August 15, 2002, 23:02:07
```

Job time: 5573 sec

```
OM of: US-09-613-508-2 to: Issued_Patents_NA:*   out_format : pfs

Date: Aug 15, 2002  8:47 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO_spool/US09613508/runat_14082002_112647_7244/app_query.fasta_1.701
 -DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US09613508_@CGN1_1_54 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-613-508-2
Query length: 303
Database: Issued_Patents_NA:*
Database sequences: 383533
Database length: 122816752
Search time (sec): 63.980000



score_list:
Sequence            Strd Orig    ZScore   EScore Len  ! Documentation  ..
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-1 + 1580.00 3605.56  3.8e-193  1371  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-308-5 +  721.00 1628.75  4.9e-83   1001  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-313-5 +  721.00 1628.75  4.9e-83   1001  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-069-138-5 +  721.00 1628.75  4.9e-83   1001  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-591-605-1 +  721.00 1627.00  6.2e-83   1158  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-3 +  721.00 1626.99  6.2e-83   1159  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-3 +  718.00 1610.21  5.3e-82   2635  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-308-9 +  711.00 1605.69  9.5e-82   1001  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-313-9 +  711.00 1605.69  9.5e-82   1001  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-069-138-9 +  711.00 1605.69  9.5e-82   1001  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-890-542A-1 +  709.00 1599.14  2.2e-81   1177  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-890-542A-3 +  709.00 1599.14  2.2e-81   1177  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-10 +  703.00 1584.14  1.5e-80   1297  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-773-608A-1 +  552.00 1227.36  1.1e-60   2658  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-522-813-1 +  468.00 1043.40  2.0e-50   1183  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-446-925-4 +  466.00 1035.59  5.4e-50   1545  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-146-331-4 +  466.00 1035.59  5.4e-50   1545  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-896-885-4 +  466.00 1035.59  5.4e-50   1545  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-375-256-4 +  466.00 1035.59  5.4e-50   1545  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-487-445-10 +  447.50  993.37  1.2e-47   1489  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-665-220-1 +  444.50  984.87  3.6e-47   1700  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-291-692-1 +  444.50  984.87  3.6e-47   1700  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-618-408B-1 +  444.50  976.07  1.1e-46   3535  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-807-200-1 +  433.50  962.45  6.4e-46   1330  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-001-777-1 +  433.50  962.45  6.4e-46   1330  !
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US96-10521-33 +  426.50  945.33  5.7e-45   1443
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-807-200-3 +  425.50  944.41  6.5e-45   1286  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-001-777-3 +  425.50  944.41  6.5e-45   1286  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-718-964-1 +  425.50  935.73  2.0e-44   2649  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-059-964A-1 +  425.50  935.73  2.0e-44   2649  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-842-341-1 +  425.50  935.73  2.0e-44   2649  !
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US96-10521-14 +  425.50  934.70  2.2e-44   2887
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-807-200-11 +  424.50  940.75  1.0e-44   1440  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-001-777-11 +  424.50  940.75  1.0e-44   1440  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-360-017-2 +  423.50  944.49  6.4e-45   870  !  S
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-618-408B-3 +  423.50  939.17  1.3e-44   1355  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-665-220-3 +  423.50  935.22  2.1e-44   1883  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-487-445-3 +  423.50  935.22  2.1e-44   1883  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-291-692-3 +  423.50  935.22  2.1e-44   1883  !
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US96-10521-17 +  423.50  931.26  3.5e-44   2619
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-852-782-2 +  413.50  915.38  2.7e-43   1440
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-852-782-1 +  413.50  913.79  3.3e-43   1644
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-446-925-6 +  396.00  876.14  4.1e-41   1313
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-146-331-6 +  396.00  876.14  4.1e-41   1313
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-896-885-6 +  396.00  876.14  4.1e-41   1313

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-1

seq_documentation_block:
; Sequence 1, Application US/08462969B
; Patent No. 6087150
;  GENERAL INFORMATION:
;    APPLICANT:  He, Wei-Wu et al.
;    TITLE OF INVENTION:  Interleukin-1 Beta Converting Enzyme
;    TITLE OF INVENTION:  Like Apoptosis Protease 3 and 4
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Ave.
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,969B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/334,251
;      FILING DATE:  11-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF140P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1371 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-462-969B-1


alignment_scores:
            Quality: 1580.00              Length:    305
              Ratio:  5.215                Gaps:      2
 Percent Similarity:  99.344   Percent Identity: 98.689

alignment_block:
 US-09-613-508-2 x US-08-462-969B-1   ..

 Align seg 1/1  to: US-08-462-969B-1  from: 1  to: 1371

       1 MetAlaAspAsp.GlnGlyCysIleGluGluGlnGlyValGluAspSerA 17
         ||||||||||||| ||||||||||||||||||||||||||||||||||||
     339 ATGGCAGATGATTCAGGGCTGTATTGAAGAGCAGGGGGTTGAGGATTCAG 388

      17 laAsnGluAspSerValAspAlaLysProAspArgSerSerPheValPro 33
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     389 CAAATGAAGATTCAGTGGATGCTAAGCCAGACCGGTCCTCGTTTGTACCG 438

      34 SerLeuPheSerLysLysLysLysAsnValThrMetArgSerIleLysTh 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     439 TCCCTCTTCAGTAAGAAGAAGAAAAATGTCACCATGCGATCCATCAAGAC 488
```

```
     50 rThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysL 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    489 CACCCGGGACCGAGTGCCTACATATCAGTACAACATGAATTTTGAAAAGC 538

     67 euGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGly 83
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    539 TGGGCAAATGCATCATAATAAACAACAAGAACTTTGATAAAGTGACAGGT 588

     84 MetGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCy 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    589 ATGGGCGTTCGAAACGGAACAGACAAAGATGCCGAGGCGCTCTTCAAGTG 638

    100 sPheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysA 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    639 CTTCCGAAGCCTGGGTTTTGACGTGATTGTCTATAATGACTGCTCTTGTG 688

    117 laLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsn 133
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    689 CCAAGATGCAAGATCTGCTTAAAAAAGCTTCTGAAGAGGACCATACAAAT 738

    134 AlaAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIl 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    739 GCCGCCTGCTTCGCCTGCATCCTCTTAAGCCATGGAGAAGAAAATGTAAT 788

    150 eTyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheA 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    789 TTATGGGAAAGATGGTGTCACACCAATAAAGGATTTGACAGCCCACTTTA 838

    167 rgGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIle 183
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    839 GGGGGGATAGATGCAAAACCCTTTTAGAGAAACCCAAACTCTTCTTCATT 888

    184 GlnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGl 200
        ||||||||||||||||||||||||||||||:::||||||||||||||||||
    889 CAGGCTTGCCGAGGGACCGAGCTTGATGATGGCATCCAGGCCGACTCGGG 938

    200 yProIleAsnAspThrAspAlaAsnProArgTyrLysIleProValGluA 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    939 GCCCATCAATGACACAGATGCTAATCCTCGATACAAGATCCCAGTGGAAG 988

    217 laAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArg 233
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    989 CTGACTTCCTCTTCGCCTATTCCACGGTTCCAGGCTATTACTCGTGGAGG 1038

    234 SerProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGl 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1039 AGCCCAGGAAGAGGCTCCTGGTTTGTGCAAGCCCTCTGCTCCATCCTGGA 1088

    250 uGluHisGlyLysGluLeuGluIleMetGlnIleLeu.ThrArgValAsn 266
        ||||||||||||||:::|||||||||||||||||||| ||||||||||||
   1089 GGAGCACGGAAAAGACCTGGAAATCATGCAAATCCTCCACCAGGGTGAAT 1138

    267 AspArgValAlaArgHisPheGluSerGlnSerAspAspProHisPheHi 283
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1139 GACAGAGTTGCCAGGCACTTTGAGTCTCAGTCTGATGACCCACACTTCCA 1188

    283 sGluLysLysGlnIleProCysValValSerMetLeuThrLysGluLeuT 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1189 TGAGAAGAAGCAGATCCCCTGTGTGGTCTCCATGCTCACCAAGGAACTCT 1238

    300 yrPheSerGln 303
        ||||||||||
   1239 ACTTCAGTCAA 1249

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-308-5

seq_documentation_block:
; Sequence 5, Application US/08964308
; Patent No. 6066715
;  GENERAL INFORMATION:
;     APPLICANT:  DESMARAIS, SYLVIE
;     APPLICANT:  FRIESEN, RICHARD
;     APPLICANT:  ZAMBONI, ROBERT
;     TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;     TITLE OF INVENTION:  BINDING ASSAY
;     NUMBER OF SEQUENCES:  15
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  ROBERT J. NORTH - MERCK & CO., INC.
;       STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;       CITY:  RAHWAY
;       STATE:  NJ
;       COUNTRY:  USA
;       ZIP:  07065
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy Diskette
;       COMPUTER:  IBM PC Compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/964,308
;       FILING DATE:  04-NOV-1996
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  NORTH, ROBERT J
;       REGISTRATION NUMBER:  27,366
;       REFERENCE/DOCKET NUMBER:  19840 PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  732-594-7262
;       TELEFAX:  732-594-4720
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1001 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-964-308-5


alignment_scores:
            Quality:  721.00              Length:    283
              Ratio:  3.353                 Gaps:      3
 Percent Similarity:  75.972   Percent Identity: 50.883

alignment_block:
 US-09-613-508-2 x US-08-964-308-5   ..

 Align seg 1/1  to: US-08-964-308-5  from: 1  to: 1001

      19 GluAspSerValAspAlaLysProAspArgSerSerPheValProSerLe 35
         |||:::|||||||||:::|||      :::   ::::::   |||:::::
      76 GAAAACTCAGTGGATTCAAAATCCATTAAA...AATTTGGAACCAAAGAT 122

      35 uPheSerLysLysLysLysAsnValThrMetArgSerIleLysThrThrA 52
         ::::      :::::::::      :::      |||:::   ::::::
     123 CATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGT. 171

      52 rgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGly 68
                               |||:::|||::::::   ::::::|||
     172 ......................TATAAAATGGATTATCCTGAGATGGGT 198

      69 LysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGl 85
           |||||||||||||||||||||||||||   |||   |||||||||
     199 TTATGTATAATAATTAATAATAAGAATTTTCATAAGAGCACTGGAATGAC 248

      85 yValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheA 102
           |||:::|||||||||   ||||||      |||   :::   ||||
     249 ATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCA 298

     102 rgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAlaLys 118
         ||:::|||   ::::::|||         ||||||   :::      :::
     299 GAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACCTGAAGAA 348

     119 MetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAl 135
```

```
      ::   :::|||::::::   :::|||:::|||||||||::::::   ::
    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAG 398

    135 aCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrG 152
        :  |||:::|||:::|||||||||||||||||||||:::::|||:::|
    399 CAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTG 448

    152 lyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheArgGly 168
        ||   :::|||   :::   :::|||   :::|||      |||||||||
    449 GAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGG 498

    169 AspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAl 185
        |||||||||:::::|||      |||||||||||||||:::|||||||||
    499 GATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGC 548

    185 aCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGlyProI 202
        ||||||||||||||||||||   :::|||:::::|||||||||
    549 CTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTG 598

    202 leAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAlaAsp 218
          :::|||      :::        :::||||||||||||||||||||
    599 ATGATGACATGGCGTGT.........CATAAAATACCAGTGGAGGCCGAC 639

    219 PheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgSerPr 235
        ||||||:::|||||||||||||:::|||||||||||||||||||||:::
    640 TTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTC 689

    235 oGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGluGluH 252
              |||||||||||||:::|||:::||||||:::::|||:::::::
    690 AAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGT 739

    252 isGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAspArg 268
        :::::   :::||||||:::|||::|||||||||||||||||   :::
    740 ATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 789

    269 ValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisGluLy 285
        ||||||   :::|||||||||   |||   |||      ||||||   ||
    790 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAA 839

    285 sLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyrPhe 301
        ||||||||||||||||:::|||||||||||||||||||||||||||||||
    840 GAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-313-5

seq_documentation_block:
; Sequence 5, Application US/08964313
; Patent No. 6114132
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  GRESSER, MICHAEL
;    APPLICANT:  KENNEDY, BRIAN
;    APPLICANT:  NICHOLSON, DONALD
;    APPLICANT:  RAMACHANDRAN, CHIDAMBARAN
;    APPLICANT:  SKOREY, KATHRYN
;    APPLICANT:  FORD-HUTCHINSON, ANTHONY
;    TITLE OF INVENTION:  PHOSPHATASE BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,313
;      FILING DATE:  04-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/030,408
;      FILING DATE:  04-NOV-1996
;      APPLICATION NUMBER:  PCT/CA97/00825
;      FILING DATE:  03 -NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19824Y
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-313-5


alignment_scores:
           Quality:  721.00              Length:    283
             Ratio:  3.353                Gaps:      3
 Percent Similarity:  75.972   Percent Identity: 50.883

alignment_block:
 US-09-613-508-2 x US-08-964-313-5   ..

 Align seg 1/1  to: US-08-964-313-5  from: 1  to: 1001

     19 GluAspSerValAspAlaLysProAspArgSerSerPheValProSerLe 35
        |||:::|||||||||:::|||      :::   ::::::   |||:::::
     76 GAAAACTCAGTGGATTCAAAATCCATTAAA...AATTTGGAACCAAAGAT 122

     35 uPheSerLysLysLysLysAsnValThrMetArgSerIleLysThrThrA 52
        ::::      ::::::::::      :::      |||:::   ::::::
    123 CATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGT. 171

     52 rgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGly 68
                                 |||:::|||::::::   ::::::|||
    172 .....TATAAAATGGATTATCCTGAGATGGGT 198

     69 LysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGl 85
        ||||||||||||||||||||||||||||||||   |||   |||||||||
    199 TTATGTATAATAATTAATAATAAGAATTTTCATAAGAGCACTGGAATGAC 248

     85 yValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheA 102
           |||:::|||||||||   ||||||      |||   :::   ||||
    249 ATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCA 298

    102 rgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAlaLys 118
        ||:::|||   :::::::|||      ||||||   :::      :::
    299 GAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

    119 MetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAl 135
        :::   :::|||::::::   :::|||:::||||||||:::::   ::
    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAG 398

    135 aCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrG 152
        :  |||:::|||:::|||||||||||||||||||||:::::|||:::|
    399 CAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTG 448

    152 lyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheArgGly 168
        ||   :::|||   :::   :::|||   :::|||      |||||||||
    449 GAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGG 498

    169 AspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAl 185
        |||||||||:::::|||      |||||||||||||||:::|||||||||
```

```
     499 GATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGC 548

     185 aCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGlyProI 202
         ||||||||||||||||||||||   :::|||:::::::|||||||||
     549 CTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTG 598

     202 leAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAlaAsp 218
           :::|||    :::          :::||||||||||||||||||||
     599 ATGATGACATGGCGTGT.........CATAAAATACCAGTGGAGGCCGAC 639

     219 PheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgSerPr 235
         ||||||:::|||||||||||||:::|||||||||||||||||||||:::
     640 TTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTC 689

     235 oGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGluGluH 252
               ||||||||||||:::|||:::||||||:::::::|||:::::::
     690 AAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGT 739

     252 isGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAspArg 268
         :::::   :::||||||:::|||:::|||||||||||||||||   :::
     740 ATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 789

     269 ValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisGluLy 285
         ||||||   :::|||||||||   |||  |||      ||||||   ||
     790 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAA 839

     285 sLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyrPhe 301
         |||||||||||||||:::|||||||||||||||||||||||||||||||
     840 GAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-069-138-5

seq_documentation_block:
; Sequence 5, Application US/09069138
; Patent No. 6348572
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  DUFRESNE, CLAUDE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  LEBLANC, YVES
;    APPLICANT:  ROY, PATRICK
;    APPLICANT:  YOUNG, ROBERT N.
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/069,138
;      FILING DATE:  29-APR-1998
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19840YIA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-069-138-5


alignment_scores:
            Quality:  721.00              Length:    283
              Ratio:  3.353                 Gaps:      3
 Percent Similarity:  75.972   Percent Identity: 50.883

alignment_block:
 US-09-613-508-2 x US-09-069-138-5   ..

 Align seq 1/1  to: US-09-069-138-5  from: 1  to: 1001

      19 GluAspSerValAspAlaLysProAspArgSerSerPheValProSerLe 35
         |||:::|||||||||:::|||      :::   ::::::   |||:::::
      76 GAAAACTCAGTGGATTCAAAATCCATTAAA...AATTTGGAACCAAAGAT 122

      35 uPheSerLysLysLysLysAsnValThrMetArgSerIleLysThrThrA 52
         ::::      :::::::::      :::      |||:::   ::::::
     123 CATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGT. 171

      52 rgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGly 68
                                 |||:::|||::::::   ::::::|||
     172 .....TATAAAATGGATTATCCTGAGATGGGT 198

      69 LysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGl 85
           ||||||||||||||||||||||||||||   |||   |||||||||
     199 TTATGTATAATAATTAATAATAAGAATTTTCATAAGAGCACTGGAATGAC 248

      85 yValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheA 102
           |||:::|||||||||   ||||||      |||   :::   ||||
     249 ATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCA 298

     102 rgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAlaLys 118
         ||:::|||   ::::::|||         ||||||   :::      :::
     299 GAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

     119 MetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAl 135
         :::   :::|||::::::   :::|||:::|||||||||:::::   ::
     349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAG 398

     135 aCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrG 152
         :   |||:::|||:::|||||||||||||||||||||:::::::|||:::|
     399 CAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTG 448

     152 lyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheArgGly 168
         ||  :::|||   :::   :::|||  :::|||      ||||||||
     449 GAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGG 498

     169 AspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAl 185
         |||||||||:::::|||      ||||||||||||||:::|||||||||
     499 GATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGC 548

     185 aCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGlyProI 202
         ||||||||||||||||||||||   :::|||:::::::|||||||||
     549 CTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTG 598

     202 leAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAlaAsp 218
           :::|||    :::          :::||||||||||||||||||||
     599 ATGATGACATGGCGTGT.........CATAAAATACCAGTGGAGGCCGAC 639

     219 PheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgSerPr 235
         ||||||:::|||||||||||||:::|||||||||||||||||||||:::
     640 TTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTC 689

     235 oGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGluGluH 252
               ||||||||||||:::|||:::||||||:::::::|||:::::::
     690 AAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGT 739
```

```
     252 isGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAspArg 268
         :::::   :::|||||||:::|||:::|||||||||||||||||||   :::
     740 ATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 789

     269 ValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisGluLy 285
         ||||||   :::||||||||||   |||   |||      ||||||   ||
     790 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAA 839

     285 sLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyrPhe 301
         |||||||||||||||:::|||||||||||||||||||||||||||||||
     840 GAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-591-605-1

seq_documentation_block:
; Sequence 1, Application US/08591605
; Patent No. 6060238
;  GENERAL INFORMATION:
;    APPLICANT:  Dixit, Vishva M.
;    TITLE OF INVENTION:  METHOD AND COMPOSITION FOR REGULATING
;    TITLE OF INVENTION:  APOPTOSIS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 PAGE MILL ROAD
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/591,605
;      FILING DATE:  09-FEB-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KONSKI, ANTOINETTE F.
;      REGISTRATION NUMBER:  34,202
;      REFERENCE/DOCKET NUMBER:  20344-21036.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  706141 MRSNFOERS SFO
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1158 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  224..1054
;      OTHER INFORMATION:  /product= "Yama peptide sequence"
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  704..718
;      OTHER INFORMATION:  /note= "region coding for
;      OTHER INFORMATION:  pentapeptide, wherein cysteine is catalytic"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "atg")
;      OTHER INFORMATION:  /note= "Mutation of catalytic cysteine to methionine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gcc")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gcg")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gct")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
;    FEATURE:
;      NAME/KEY:  mutation
;      LOCATION:  replace(710..712, "gca")
;      OTHER INFORMATION:  /note= "mutation of catalytic cysteine to alanine"
US-08-591-605-1


alignment_scores:
            Quality:  721.00              Length:    283
              Ratio:  3.353                Gaps:      3
 Percent Similarity:  75.972   Percent Identity: 50.883

alignment_block:
 US-09-613-508-2 x US-08-591-605-1   ..

 Align seg 1/1  to: US-08-591-605-1  from: 1  to: 1158

      19 GluAspSerValAspAlaLysProAspArgSerSerPheValProSerLe 35
         |||::|||||||||:::|||      :::   ::::::   |||:::::
     236 GAAAACTCAGTGGATTCAAAATCCATTAAA...AATTTGGAACCAAAGAT 282

      35 uPheSerLysLysLysLysAsnValThrMetArgSerIleLysThrThrA 52
         ::::      :::::::::      :::      |||:::   ::::::
     283 CATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGT. 331

      52 rgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGly 68
                                 |||:::|||::::::   ::::::|||
     332 .....TATAAAATGGATTATCCTGAGATGGGT 358

      69 LysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGl 85
         ||||||||||||||||||||||||||||||   |||   |||||||||
     359 TTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGAC 408

      85 yValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheA 102
            |||:::|||||||||   ||||||      |||   :::   ||||
     409 ATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCA 458

     102 rgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAlaLys 118
         ||:::|||   ::::::|||          ||||||   :::      :::
     459 GAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 508

     119 MetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAl 135
         :::   :::|||:::::::   :::|||:::|||||||||:::::   ::
     509 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAG 558

     135 aCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrG 152
         :   |||:::|||:::|||||||||||||||||||||:::::|||:::|
     559 CAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTG 608

     152 lyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheArgGly 168
         ||   :::|||   :::   :::|||   :::|||      |||||||||
     609 GAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGG 658

     169 AspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAl 185
         ||||||||||:::::|||      |||||||||||||||:::|||||||||
     659 GATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGC 708

     185 aCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGlyProI 202
         ||||||||||||||||||||   :::|||:::::|||||||||
     709 CTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTG 758

     202 leAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAlaAsp 218
           :::|||      :::         :::|||||||||||||||||||||
     759 ATGATGACATGGCGTGT.........CATAAAATACCAGTGGAGGCCGAC 799

     219 PheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgSerPr 235
         ||||||:::|||||||||||||:::|||||||||||||||||||||||:::
     800 TTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTC 849
```

```
     235 oGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGluGluH 252
     850 AAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGT 899

     252 isGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAspArg 268
     900 ATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 949

     269 ValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisGluLy 285
     950 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAA 999

     285 sLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyrPhe 301
    1000 GAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1048

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-462-969B-3

seq_documentation_block:
; Sequence 3, Application US/08462969B
; Patent No. 6087150
;  GENERAL INFORMATION:
;    APPLICANT:  He, Wei-Wu et al.
;    TITLE OF INVENTION:  Interleukin-1 Beta Converting Enzyme
;    TITLE OF INVENTION:  Like Apoptosis Protease 3 and 4
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Ave.
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,969B
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/334,251
;      FILING DATE:  11-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF140P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1159 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-462-969B-3


alignment_scores:
            Quality:  721.00            Length:    283
              Ratio:  3.353               Gaps:      3
 Percent Similarity:  75.972   Percent Identity: 50.883

alignment_block:
 US-09-613-508-2 x US-08-462-969B-3   ..

 Align seg 1/1  to: US-08-462-969B-3  from: 1  to: 1159

      19 GluAspSerValAspAlaLysProAspArgSerSerPheValProSerLe 35

     235 GAAAACTCAGTGGATTCAAAATCCATTAAA...AATTTGGAACCAAAGAT 281

      35 uPheSerLysLysLysLysAsnValThrMetArgSerIleLysThrThrA 52
     282 CATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGT. 330

      52 rgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGly 68
     331 .....TATAAAATGGATTATCCTGAGATGGGT 357

      69 LysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGl 85
     358 TTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGAC 407

      85 yValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheA 102
     408 ATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCA 457

     102 rgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAlaLys 118
     458 GAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 507

     119 MetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAl 135
     508 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAG 557

     135 aCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrG 152
     558 CAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTG 607

     152 lyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheArgGly 168
     608 GAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGG 657

     169 AspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAl 185
     658 GATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGC 707

     185 aCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGlyProI 202
     708 CTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTG 757

     202 leAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAlaAsp 218
     758 ATGATGACATGGCGTGT.........CATAAAATACCAGTGGAGGCCGAC 798

     219 PheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgSerPr 235
     799 TTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTC 848

     235 oGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGluGluH 252
     849 AAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGT 898

     252 isGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAspArg 268
     899 ATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 948

     269 ValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisGluLy 285
     949 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAA 998

     285 sLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyrPhe 301
     999 GAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1047

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-3

seq_documentation_block:
; Sequence 3, Application US/09484617
; Patent No. 6303374
; GENERAL INFORMATION:
```

```
;  APPLICANT: Hong Zhang
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF CASPASE 3 EXPRESSION
;  FILE REFERENCE: RTS-0103
;  CURRENT APPLICATION NUMBER: US/09/484,617
;  CURRENT FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 176
; SEQ ID NO 3
;   LENGTH: 2635
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (225)...(1058)
US-09-484-617-3


alignment_scores:
            Quality:  718.00              Length:    283
              Ratio:  3.340                Gaps:      3
 Percent Similarity:  75.972   Percent Identity: 50.530

alignment_block:
 US-09-613-508-2 x US-09-484-617-3   ..

 Align seg 1/1  to: US-09-484-617-3  from: 1  to: 2635

      19 GluAspSerValAspAlaLysProAspArgSerSerPheValProSerLe 35
         |||:::||||||||||:::|||      :::   ::::::   |||::::::
     237 GAAAACTCAGTGGATTCAAAATCCATTAAA...AATTTGGAACCAAAGAT 283

      35 uPheSerLysLysLysLysAsnValThrMetArgSerIleLysThrThrA 52
         ::::      :::::::::      :::      |||:::   ::::::
     284 CATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGT. 332

      52 rgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGly 68
                               |||:::|||::::::   ::::::|||
     333 .....TATAAAATGGATTATCCTGAGATGGGT 359

      69 LysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGl 85
         |||||||||||||||||||||||||||||   |||   |||||||||
     360 TTATGTATAATAATTAATAATAAGAATTTTCATAAAAGCACTGGAATGAC 409

      85 yValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheA 102
         |||:::|||||||||   ||||||      |||   :::   ||||
     410 ATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCA 459

     102 rgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAlaLys 118
         ||:::|||   ::::::|||         ||||||   :::      :::
     460 GAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 509

     119 MetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAl 135
         :::   :::|||::::::   :::|||:::||||||||:::::   ::
     510 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAG 559

     135 aCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrG 152
         :   |||:::|||:::||||||||||||||||||||:::::|||:::|
     560 CAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTG 609

     152 lyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheArgGly 168
         ||   :::|||   :::   :::|||   :::|||      |||||||||
     610 GAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGG 659

     169 AspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAl 185
         |||||||||:::::::|||      ||||||||||||||:::|||||||
     660 GATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGC 709

     185 aCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGlyProI 202
         ||||||||||||||||||||   :::|||::::::|||||||||
     710 CTGCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTG 759

     202 leAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAlaAsp 218
           :::|||      :::         :::||||||||||||:::||||||
     760 ATGATGACATGGCGTGT.........CATAAAATACCAGTGGATGCCGAC 800

     219 PheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgSerPr 235
         ||||||:::|||||||||||||:::|||||||||||||||||||||:::
     801 TTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTC 850

     235 oGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGluGluH 252
             ||||||||||||:::|||:::||||||:::::::|||:::::::
     851 AAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGT 900

     252 isGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAspArg 268
         :::::   :::||||||:::|||:::|||||||||||||||||   :::
     901 ATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 950

     269 ValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisGluLy 285
         ||||||   :::|||||||||   |||   |||      ||||||   ||
     951 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAA 1000

     285 sLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyrPhe 301
         |||||||||||||||:::|||||||||||||||||||||||||||||||
    1001 GAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 1049

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-308-9

seq_documentation_block:
; Sequence 9, Application US/08964308
; Patent No. 6066715
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBERT J. NORTH - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,308
;      FILING DATE:  04-NOV-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  NORTH, ROBERT J
;      REGISTRATION NUMBER:  27,366
;      REFERENCE/DOCKET NUMBER:  19840 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-7262
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-308-9


alignment_scores:
            Quality:  711.00              Length:    283
              Ratio:  3.322                Gaps:      3
 Percent Similarity:  75.618   Percent Identity: 50.530
```

```
alignment_block:
 US-09-613-508-2 x US-08-964-308-9   ..

 Align seg 1/1  to: US-08-964-308-9  from: 1  to: 1001

      19 GluAspSerValAspAlaLysProAspArgSerSerPheValProSerLe 35
      76 GAAAACTCAGTGGATTCAAAATCCATTAAA...AATTTGGAACCAAAGAT 122

      35 uPheSerLysLysLysLysAsnValThrMetArgSerIleLysThrThrA 52
     123 CATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGT. 171

      52 rgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGly 68
     172 .....TATAAAATGGATTATCCTGAGATGGGT 198

      69 LysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGl 85
     199 TTATGTATAATAATTAATAATAAGAATTTTCATAAGAGCACTGGAATGAC 248

      85 yValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheA 102
     249 ATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCA 298

     102 rgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAlaLys 118
     299 GAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

     119 MetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAl 135
     349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAG 398

     135 aCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrG 152
     399 CAGTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTG 448

     152 lyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheArgGly 168
     449 GAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGG 498

     169 AspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAl 185
     499 GATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGC 548

     185 aCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGlyProI 202
     549 CTCCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTG 598

     202 leAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAlaAsp 218
     599 ATGATGACATGGCGTGT.........CATAAAATACCAGTGGAGGCCGAC 639

     219 PheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgSerPr 235
     640 TTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTC 689

     235 oGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGluGluH 252
     690 AAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGT 739

     252 isGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAspArg 268
     740 ATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 789

     269 ValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisGluLy 285
     790 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAA 839

     285 sLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyrPhe 301
     840 GAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888
```

```
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-964-313-9

seq_documentation_block:
; Sequence 9, Application US/08964313
; Patent No. 6114132
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  GRESSER, MICHAEL
;    APPLICANT:  KENNEDY, BRIAN
;    APPLICANT:  NICHOLSON, DONALD
;    APPLICANT:  RAMACHANDRAN, CHIDAMBARAN
;    APPLICANT:  SKOREY, KATHRYN
;    APPLICANT:  FORD-HUTCHINSON, ANTHONY
;    TITLE OF INVENTION:  PHOSPHATASE BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/964,313
;      FILING DATE:  04-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/030,408
;      FILING DATE:  04-NOV-1996
;      APPLICATION NUMBER:  PCT/CA97/00825
;      FILING DATE:  03 -NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19824Y
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-964-313-9


alignment_scores:
            Quality:  711.00              Length:    283
              Ratio:  3.322                 Gaps:      3
 Percent Similarity:  75.618   Percent Identity: 50.530

alignment_block:
 US-09-613-508-2 x US-08-964-313-9   ..

 Align seg 1/1  to: US-08-964-313-9  from: 1  to: 1001

      19 GluAspSerValAspAlaLysProAspArgSerSerPheValProSerLe 35
      76 GAAAACTCAGTGGATTCAAAATCCATTAAA...AATTTGGAACCAAAGAT 122

      35 uPheSerLysLysLysLysAsnValThrMetArgSerIleLysThrThrA 52
     123 CATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGT. 171
```

```
     52 rgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGly 68
                                 |||:::|||:::::   ::::::|||
    172 .....................TATAAAATGGATTATCCTGAGATGGGT 198

     69 LysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGl 85
        |||||||||||||||||||||||||||  |||   |||||||||
    199 TTATGTATAATAATTAATAATAAGAATTTTCATAAGAGCACTGGAATGAC 248

     85 yValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheA 102
          |||:::|||||||||   ||||||      |||   :::   ||||
    249 ATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCA 298

    102 rgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAlaLys 118
        ||:::|||   :::::::|||         ||||||   :::      :::
    299 GAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

    119 MetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAl 135
        :::   :::|||::::::   :::|||:::|||||||||:::::   ::
    349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAG 398

    135 aCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrG 152
        :   |||:::|||:::||||||||||||||||||:::::|||:::|
    399 CAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTG 448

    152 lyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheArgGly 168
        ||   :::|||   :::   :::|||   :::|||      |||||||||
    449 GAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGG 498

    169 AspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAl 185
        |||||||||:::::::|||      |||||||||||||||:::|||||||
    499 GATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGC 548

    185 aCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGlyProI 202
        |   ||||||||||||||||||   :::|||:::::|||||||||
    549 CTCCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTG 598

    202 leAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAlaAsp 218
          :::|||      :::         :::||||||||||||||||||||
    599 ATGATGACATGGCGTGT.........CATAAAATACCAGTGGAGGCCGAC 639

    219 PheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgSerPr 235
        ||||||:::|||||||||||||:::|||||||||||||||||||||||:::
    640 TTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTC 689

    235 oGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGluGluH 252
            ||||||||||||:::|||:::||||||:::::::|||:::::::
    690 AAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGT 739

    252 isGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAspArg 268
        :::::   :::|||||||:::|||:::|||||||||||||||||   :::
    740 ATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 789

    269 ValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisGluLy 285
        ||||||   :::||||||||   |||   |||      ||||||   ||
    790 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAA 839

    285 sLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyrPhe 301
        |||||||||||||||:::|||||||||||||||||||||||||||||||
    840 GAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-069-138-9

seq_documentation_block:
; Sequence 9, Application US/09069138
; Patent No. 6348572
;  GENERAL INFORMATION:
;    APPLICANT:  DESMARAIS, SYLVIE
;    APPLICANT:  DUFRESNE, CLAUDE
;    APPLICANT:  FRIESEN, RICHARD
;    APPLICANT:  LEBLANC, YVES
;    APPLICANT:  ROY, PATRICK
;    APPLICANT:  YOUNG, ROBERT N.
;    APPLICANT:  ZAMBONI, ROBERT
;    TITLE OF INVENTION:  NEW LIGANDS FOR PHOSPHATASE
;    TITLE OF INVENTION:  BINDING ASSAY
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILIPPE L. DURETTE - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Diskette
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/069,138
;      FILING DATE:  29-APR-1998
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DURETTE, PHILIPPE L.
;      REGISTRATION NUMBER:  35,125
;      REFERENCE/DOCKET NUMBER:  19840YIA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-4568
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1001 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-069-138-9


alignment_scores:
            Quality:  711.00              Length:    283
              Ratio:  3.322                Gaps:      3
 Percent Similarity:  75.618   Percent Identity: 50.530

alignment_block:
 US-09-613-508-2 x US-09-069-138-9   ..

 Align seg 1/1  to: US-09-069-138-9  from: 1  to: 1001

      19 GluAspSerValAspAlaLysProAspArgSerSerPheValProSerLe 35
         |||:::|||||||||:::|||      :::   ::::::   |||:::::
      76 GAAAACTCAGTGGATTCAAAATCCATTAAA...AATTTGGAACCAAAGAT 122

      35 uPheSerLysLysLysLysAsnValThrMetArgSerIleLysThrThrA 52
         ::::      :::::::::      :::      |||:::   ::::::
     123 CATACATGGAAGCGAATCAATGGACTCTGGAATATCCCTGGACAACAGT. 171

      52 rgAspArgValProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGly 68
                                 |||:::|||::::::   :::::::|||
     172 .....................TATAAAATGGATTATCCTGAGATGGGT 198

      69 LysCysIleIleIleAsnAsnLysAsnPheAspLysValThrGlyMetGl 85
         |||||||||||||||||||||||||||  |||   |||||||||
     199 TTATGTATAATAATTAATAATAAGAATTTTCATAAGAGCACTGGAATGAC 248

      85 yValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheLysCysPheA 102
           |||:::|||||||||   ||||||      |||   :::   ||||
     249 ATCTCGGTCTGGTACAGATGTCGATGCAGCAAACCTCAGGGAAACATTCA 298

     102 rgSerLeuGlyPheAspValIleValTyrAsnAspCysSerCysAlaLys 118
         ||:::|||   :::::|||         ||||||   :::      :::
     299 GAAACTTGAAATATGAAGTCAGGAATAAAAATGATCTTACACGTGAAGAA 348

     119 MetGlnAspLeuLeuLysLysAlaSerGluGluAspHisThrAsnAlaAl 135
         :::   :::|||::::::   :::|||:::|||||||||:::::   ::
```

```
     349 ATTGTGGAATTGATGCGTGATGTTTCTAAAGAAGATCACAGCAAAAGGAG 398

     135 aCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnValIleTyrG 152
         :  |||:::|||:::|||||||||||||||||||||||:::::::|||:::|
     399 CAGTTTTGTTTGTGTGCTTCTGAGCCATGGTGAAGAAGGAATAATTTTTG 448

     152 lyLysAspGlyValThrProIleLysAspLeuThrAlaHisPheArgGly 168
         ||   :::|||   :::   :::|||   :::|||      |||||||||
     449 GAACAAATGGACCTGTTGACCTGAAAAAAATAACAAACTTTTTCAGAGGG 498

     169 AspArgCysLysThrLeuLeuGluLysProLysLeuPhePheIleGlnAl 185
         |||||||||:::::|||      |||||||||||||||:::|||||||||
     499 GATCGTTGTAGAAGTCTAACTGGAAAACCCAAACTTTTCATTATTCAGGC 548

     185 aCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspSerGlyProI 202
         |  ||||||||||||||||||   :::|||:::::::|||||||||
     549 CTCCCGTGGTACAGAACTGGACTGTGGCATTGAGACAGACAGTGGTGTTG 598

     202 leAsnAspThrAspAlaAsnProArgTyrLysIleProValGluAlaAsp 218
           :::|||      :::         :::|||||||||||||||||||||
     599 ATGATGACATGGCGTGT.........CATAAAATACCAGTGGAGGCCGAC 639

     219 PheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgSerPr 235
         ||||||:::|||||||||||||:::|||||||||||||||||||||:::
     640 TTCTTGTATGCATACTCCACAGCACCTGGTTATTATTCTTGGCGAAATTC 689

     235 oGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGluGluH 252
             ||||||||||||:::|||:::|||||:::::::|||:::::::
     690 AAAGGATGGCTCCTGGTTCATCCAGTCGCTTTGTGCCATGCTGAAACAGT 739

     252 isGlyLysGluLeuGluIleMetGlnIleLeuThrArgValAsnAspArg 268
         :::::   :::|||||||:::|||:::|||||||||||||||||   :::
     740 ATGCCGACAAGCTTGAATTTATGCACATTCTTACCCGGGTTAACCGAAAG 789

     269 ValAlaArgHisPheGluSerGlnSerAspAspProHisPheHisGluLy 285
         ||||||   :::|||||||||   |||   |||      ||||||   ||
     790 GTGGCAACAGAATTTGAGTCCTTTTCCTTTGACGCTACTTTTCATGCAAA 839

     285 sLysGlnIleProCysValValSerMetLeuThrLysGluLeuTyrPhe 301
         ||||||||||||||||:::|||||||||||||||||||||||||||||
     840 GAAACAGATTCCATGTATTGTTTCCATGCTCACAAAAGAACTCTATTTT 888

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-890-542A-1

seq_documentation_block:
; Sequence 1, Application US/08890542A
; Patent No. 5840509
;  GENERAL INFORMATION:
;    APPLICANT:  Ni, Binhui
;    APPLICANT:  Paul, Steven M
;    APPLICANT:  Wu, Xin
;    TITLE OF INVENTION:  PROTEASE AND RELATED NUCLEIC ACID
;    TITLE OF INVENTION:  COMPOUNDS
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  United States of America
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/890,542A
;      FILING DATE:  09-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gaylo, Paul J.
;      REGISTRATION NUMBER:  36,808
;      REFERENCE/DOCKET NUMBER:  X-10704
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (317) 276-0756
;      TELEFAX:  (317) 276-3861
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1177 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  76..906
US-08-890-542A-1


alignment_scores:
           Quality:  709.00              Length:    286
             Ratio:  3.298                 Gaps:      3
 Percent Similarity:  75.175   Percent Identity: 50.000

alignment_block:
 US-09-613-508-2 x US-08-890-542A-1   ..

 Align seg 1/1  to: US-08-890-542A-1  from: 1  to: 1177

      18 AsnGluAspSerValAspAlaLysProAspArgSerSerPheValProSe 34
         ||||||   |||||||||:::|||      ::::::      :::   ::
      85 AACGAAACCTCCGTGGATTCAAAATCCATTAATAATTTTGAAACAAAGAC 134

      34 rLeuPheSerLysLysLys......LysAsnValThrMetArgSerIleL 49
         ::::   ::::::|||:::      :::::::::   :::   |||
     135 TATCCATGGAAGCAAGTCGATGGACTCTGGAATATATCTGGACAGC.... 180

      49 ysThrThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGlu 65
                                        :::|||:::|||::::::
     181 ..........................AGTTACAAAATGGATTACCCT 201

      66 LysLeuGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValTh 82
         ::::::|||   |||||||||||||||||||||||||   |||   ||
     202 GAAATGGGCTTGTGTATAATAATTAATAATAAGAACTTCCATAAAAGCAC 251

      82 rGlyMetGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheL 99
         |||||||::::::|||||||||||||||   ||||||      |||   :
     252 TGGAATGTCAGCTCGCAATGGTACCGATGTCGATGCAGCTAACCTCAGAG 301

      99 ysCysPheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSer 115
         ::   |||   :::|||   ::::::|||          |||||   :::
     302 AGACATTCATGGCCCTGAAATACGAAGTCAGGAATAAAAATGACCTTACT 351

     116 CysAlaLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisTh 132
             :::::::::::::|||:::   ::::::|||:::||||||||||::
     352 CGTGAAGAAATTATGGAATTGATGGATAGTGTTTCTAAGGAAGATCACAG 401

     132 rAsnAlaAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnV 149
         ::::   :::   |||:::|||::::::||||||||||||:::|||:::|
     402 CAAAAGGAGCAGTTTTGTGTGTGTGATTCTAAGTCATGGAGATGAAGGAG 451

     149 alIleTyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHis 165
         |||||:::|||   :::|||   :::   :::|||   |||||||:::
     452 TAATTTTTGGAACGAACGGACCTGTGGACCTGAAAAAACTAACTAGTTTC 501

     166 PheArgGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePh 182
         ||||||||||||   |||::::::|||       |||||||||||||::
     502 TTCAGAGGCGACTACTGCCGGAGTCTGACTGGAAAGCCGAAACTCTTCAT 551

     182 eIleGlnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspS 199
         :|||||||||||||||||||||||||||||||   :::|||:::::::||||
     552 CATTCAGGCCTGCCGAGGTACAGAGCTGGACTGCGGTATTGAGACAGACA 601

     199 erGlyProIleAsnAspThrAspAlaAsnProArgTyrLysIleProVal 215
         |||||      :::|||      ::::::         ||||||||||||
```

```
     602 GTGGAACTGACGATGATATGGCATGCCAG.........AAGATACCAGTG 642

     216 GluAlaAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTr 232
         ||||||||||||:::|||||||||||||:::|||||||||||||||||
     643 GGGGCCGACTTCCTGTATGCTTACTCTACCGCACCCGGTTACTATTCCTG 692

     232 pArgSerProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleL 249
         ||||:::          ||||||||||||:::|||:::|||||:::::|
     693 GAGAAATTCAAGGGACGGGTCATGGTTCATCCAGTCACTTTGCGCCATGC 742

     249 euGluGluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgVal 265
         ||:::   ::::::   :::||||||:::|||:::|||||||||||||||
     743 TGAAACTGTACGCGCACAAGCTGGAATTCATGCACATCCTCACTCGTGTT 792

     266 AsnAspArgValAlaArgHisPheGluSerGlnSerAspAspProHisPh 282
         |||   :::||||||   :::|||||||||   |||   |||      ||
     793 AACCGGAAGGTGGCCATGGAATTTGAGTCCTTCTCCCTGGACGCCACTTT 842

     282 eHisGluLysLysGlnIleProCysValValSerMetLeuThrLysGluL 299
         ||||  ||||||||||||||||||:::||||||||||||||||||||||
     843 CCACGCAAAGAAACAGATCCCGTGTATTGTGTCAATGCTCACAAAAGAAC 892

     299 euTyrPhe 301
         ||||||||
     893 TGTACTTT 900

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-890-542A-3

seq_documentation_block:
; Sequence 3, Application US/08890542A
; Patent No. 5840509
;  GENERAL INFORMATION:
;    APPLICANT:  Ni, Binhui
;    APPLICANT:  Paul, Steven M
;    APPLICANT:  Wu, Xin
;    TITLE OF INVENTION:  PROTEASE AND RELATED NUCLEIC ACID
;    TITLE OF INVENTION:  COMPOUNDS
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  United States of America
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/890,542A
;      FILING DATE:  09-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gaylo, Paul J.
;      REGISTRATION NUMBER:  36,808
;      REFERENCE/DOCKET NUMBER:  X-10704
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (317) 276-0756
;      TELEFAX:  (317) 276-3861
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1177 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  mRNA
US-08-890-542A-3


alignment_scores:
           Quality:  709.00              Length:    286
               Ratio:  3.298                 Gaps:      3
 Percent Similarity:  75.175   Percent Identity: 50.000

alignment_block:
 US-09-613-508-2 x US-08-890-542A-3   ..

 Align seg 1/1  to: US-08-890-542A-3  from: 1  to: 1177

      18 AsnGluAspSerValAspAlaLysProAspArgSerSerPheValProSe 34
         ||||||   |||||||||:::|||      ::::::      :::   ::
      85 AACGAAACCUCCGUGGAUUCAAAAUCCAUUAAUAAUUUUGAAACAAAGAC 134

      34 rLeuPheSerLysLysLys......LysAsnValThrMetArgSerIleL 49
         ::::   ::::::|||:::      :::::::::   :::   |||
     135 UAUCCAUGGAAGCAAGUCGAUGGACUCUGGAAUAUAUCUGGACAGC.... 180

      49 ysThrThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGlu 65
                                      :::|||:::|||::::::
     181 ..........................AGUUACAAAAUGGAUUACCCU 201

      66 LysLeuGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValTh 82
         ::::::|||   ||||||||||||||||||||||||||   |||   ||
     202 GAAAUGGGCUUGUGUAUAAUAAUUAAUAAUAAGAACUUCCAUAAAAGCAC 251

      82 rGlyMetGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheL 99
         |||||||::::::|||||||||||||||||   ||||||      |||   :
     252 UGGAAUGUCAGCUCGCAAUGGUACCGAUGUCGAUGCAGCUAACCUCAGAG 301

      99 ysCysPheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSer 115
         ::   |||   :::|||   ::::::|||          ||||||   :::
     302 AGACAUUCAUGGCCCUGAAAUACGAAGUCAGGAAUAAAAAUGACCUUACU 351

     116 CysAlaLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisTh 132
         ::::::::::::|||:::   ::::::|||:::||||||||||||::
     352 CGUGAAGAAAUUAUGGAAUUGAUGGAUAGUGUUUCUAAGGAAGAUCACAG 401

     132 rAsnAlaAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnV 149
         . ::::   :::   |||:::|||::::::|||||||||||:::|||:::|
     402 CAAAAGGAGCAGUUUUGUGUGUGUGAUUCUAAGUCAUGGAGAUGAAGGAG 451

     149 alIleTyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHis 165
         |||||:::|||   :::|||   :::   :::|||   |||||||:::
     452 UAAUUUUUGGAACGAACGGACCUGUGGACCUGAAAAAACUAACUAGUUUC 501

     166 PheArgGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePh 182
         ||||||||||||   |||:::::|||       ||||||||||||||::
     502 UUCAGAGGCGACUACUGCCGGAGUCUGACUGGAAAGCCGAAACUCUUCAU 551

     182 eIleGlnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspS 199
         :|||||||||||||||||||||||||||||||   :::|||:::::::|||
     552 CAUUCAGGCCUGCCGAGGUACAGAGCUGGACUGCGGUAUUGAGACAGACA 601

     199 erGlyProIleAsnAspThrAspAlaAsnProArgTyrLysIleProVal 215
         |||||       :::|||       ::::::         ||||||||||||
     602 GUGGAACUGACGAUGAUAUGGCAUGCCAG.........AAGAUACCAGUG 642

     216 GluAlaAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSerTr 232
         ||||||||||||:::||||||||||||:::|||||||||||||||||
     643 GGGGCCGACUUCCUGUAUGCUUACUCUACCGCACCCGGUUACUAUUCCUG 692

     232 pArgSerProGlyArgGlySerTrpPheValGlnAlaLeuCysSerIleL 249
         ||||:::          ||||||||||||:::|||:::|||||:::::|
     693 GAGAAAUUCAAGGGACGGGUCAUGGUUCAUCCAGUCACUUUGCGCCAUGC 742

     249 euGluGluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArgVal 265
         ||:::   ::::::   :::||||||:::|||:::|||||||||||||||
     743 UGAAACUGUACGCGCACAAGCUGGAAUUCAUGCACAUCCUCACUCGUGUU 792

     266 AsnAspArgValAlaArgHisPheGluSerGlnSerAspAspProHisPh 282
         |||   :::||||||   :::|||||||||   |||   |||      ||
     793 AACCGGAAGGUGGCCAUGGAAUUUGAGUCCUUCUCCCUGGACGCCACUUU 842

     282 eHisGluLysLysGlnIleProCysValValSerMetLeuThrLysGluL 299
```

```
      |||||   |||||||||||||||||||:::|||||||||||||||||||||||
    843 CCACGCAAAGAAACAGAUCCCGUGUAUUGUGUCAAUGCUCACAAAAGAAC 892

    299 euTyrPhe 301
        ||||||||
    893 UGUACUUU 900

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-484-617-10

seq_documentation_block:
; Sequence 10, Application US/09484617
; Patent No. 6303374
; GENERAL INFORMATION:
;  APPLICANT: Hong Zhang
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF CASPASE 3 EXPRESSION
;  FILE REFERENCE: RTS-0103
;  CURRENT APPLICATION NUMBER: US/09/484,617
;  CURRENT FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 176
; SEQ ID NO 10
;   LENGTH: 1297
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (68)...(901)
US-09-484-617-10


alignment_scores:
           Quality:  703.00            Length:    287
             Ratio:  3.316              Gaps:      4
 Percent Similarity:  73.868   Percent Identity: 50.871

alignment_block:
 US-09-613-508-2 x US-09-484-617-10   ..

 Align seg 1/1  to: US-09-484-617-10  from: 1  to: 1297

     18 AsnGluAspSerValAspAlaLysProAspArgSerSerPheValProSe 34
        |||:::   |||||||||:::|||      ::::::      |||   ::
     77 AACAAAACCTCAGTGGATTCAAAATCCATTAATAATTTTGAAGTAAAGAC 126

     34 rLeuPheSerLysLysLys......LysAsnValThrMetArgSerIleL 49
        ::::   ::::::|||:::      :::::::::   :::   |||
    127 CATACATGGGAGCAAGTCAGTGGACTCTGGGATCTATCTGGACAGT.... 172

     49 ysThrThrArgAspArgValProThrTyrGlnTyrAsnMetAsnPheGlu 65
                                  :::|||:::|||::::::
    173 ............................AGTTACAAAATGGATTATCCT 193

     66 LysLeuGlyLysCysIleIleIleAsnAsnLysAsnPheAspLysValTh 82
        :::::::|||   |||||||||||||||||||||||||||||   |||   ||
    194 GAAATGGGCATATGCATAATAATTAATAATAAGAACTTCCATAAGAGCAC 243

     82 rGlyMetGlyValArgAsnGlyThrAspLysAspAlaGluAlaLeuPheL 99
        |||||||:::   |||:::|||||||||   ||||||      |||   :
    244 TGGAATGTCATCTCGCTCTGGTACGGATGTGGACGCAGCCAACCTCAGAG 293

     99 ysCysPheArgSerLeuGlyPheAspValIleValTyrAsnAspCysSer 115
        ::   |||   :::|||   :::::::|||           ||||||   :::
    294 AGACATTCATGGGCCTGAAATACCAAGTCAGGAATAAAAATGATCTTACT 343

    116 CysAlaLysMetGlnAspLeuLeuLysLysAlaSerGluGluAspHisTh 132
                :::   :::|||:::   ::::::|||:::||||||||||::
    344 CGTGAAGACATTTTGGAATTAATGGATAGTGTTTCTAAGGAAGATCATAG 393

    132 rAsnAlaAlaCysPheAlaCysIleLeuLeuSerHisGlyGluGluAsnV 149
        ::::   :::   |||:::||::::::|||||||||||||:::|||:::|
    394 CAAAAGGAGCAGCTTTGTGTGTGTGATTCTAAGCCATGGTGATGAAGGGG 443

    149 alIleTyrGlyLysAspGlyValThrProIleLysAspLeuThrAlaHis 165

        ||||||||||   :::|||   :::   :::|||   ||||||:::
    444 TCATTTATGGGACAAATGGGCCTGTTGAACTGAAAAAGTTGACTAGCTTC 493

    166 PheArgGlyAspArgCysLysThrLeuLeuGluLysProLysLeuPhePh 182
        ||||||||||||   |||:::::::|||       ||||||||||||||::
    494 TTCAGAGGCGACTACTGCCGGAGTCTGACTGGAAAGCCGAAACTCTTCAT 543

    182 eIleGlnAlaCysArgGlyThrGluLeuAspAspAlaIleGlnAlaAspS 199
        :|||||||||||||||||||||||||||||||   :::|||:::::|||||
    544 CATTCAGGCCTGCCGGGGTACGGAGCTGGACTGTGGCATTGAGACAGACA 593

    199 erGlyProIleAsnAspThrAspAlaAsnProArg...TyrLysIlePro 214
        |||||             ||||||   :::            |||||||||
    594 GTGGG.............ACTGATGAGGAGATGGCTTGCCAGAAGATACCG 631

    215 ValGluAlaAspPheLeuPheAlaTyrSerThrValProGlyTyrTyrSe 231
        ||||||||||||||||||:::||||||||||||:::|||||||||||||||
    632 GTGGAGGCTGACTTCCTGTATGCTTACTCTACAGCACCTGGTTACTATTC 681

    231 rTrpArgSerProGlyArgGlySerTrpPheValGlnAlaLeuCysSerI 248
        |||||||:::          ||||||||||||:::|||:::|||||||||:
    682 CTGGAGAAATTCAAAGGACGGGTCGTGGTTCATCCAGTCCCTTGCAGCA 731

    248 leLeuGluGluHisGlyLysGluLeuGluIleMetGlnIleLeuThrArg 264
        ::|||:::   ::::::   :::||||||:::|||:::|||||||||||||
    732 TGCTGAAGCTGTACGCGCACAAGCTAGAATTTATGCACATTCTCACTCGC 781

    265 ValAsnAspArgValAlaArgHisPheGluSerGlnSerAspAspProHi 281
        ||||||   :::||||||   :::|||||||||   |||   |||
    782 GTTAACAGGAAGGTGGCAACGGAATTCGAGTCCTTCTCCCTGGACTCCAC 831

    281 sPheHisGluLysLysGlnIleProCysValValSerMetLeuThrLysG 298
        ||||||   ||||||||||||||||||::|||||||||||||||||||||
    832 TTTCCACGCAAAGAAACAGATCCCGTGTATTGTGTCCATGCTCACGAAAG 881

    298 luLeuTyrPhe 301
        ||||||||||||
    882 AACTGTACTTT 892

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-773-608A-1

seq_documentation_block:
; Sequence 1, Application US/08773608A
; Patent No. 5858778
;  GENERAL INFORMATION:
;    APPLICANT:  Alnemri, Emad S.
;    APPLICANT:  Fernandes-Alnemri, Teresa
;    APPLICANT:  Litwack, Gerald
;    TITLE OF INVENTION:  Sf Caspase-1, Compositions and Methods
;    TITLE OF INVENTION:  for Making and Methods of Using the Same
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5858778ris
;      STREET:  One Liberty Place 46th floor
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  US
;      ZIP:  19355
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/773,608A
;      FILING DATE:  27-DEC-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeLuca, Mark
;      REGISTRATION NUMBER:  33,229
;      REFERENCE/DOCKET NUMBER:  TJU-2087
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
```

```
;       TELEFAX:  215-568-3439
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2658 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  not relevant
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  140..1039
US-08-773-608A-1


alignment_scores:
            Quality:  552.00              Length:    301
              Ratio:  2.629                Gaps:      7
 Percent Similarity:  69.767    Percent Identity: 40.532

alignment_block:
 US-09-613-508-2 x US-08-773-608A-1   ..

 Align seg 1/1  to: US-08-773-608A-1  from: 1  to: 2658

      12 GlyValGluAspSerAlaAsnGluAspSerValAspAlaLysPro..... 26
         |||   :::|||:::::::|||   ||||||||||||   |||
     149 GGAAAACAAGACAATGGAAATGTGGATAGTGTTGATATCAAACAAAGAAC 198

      27 .............AspArgSerSerPheValProSerLeuPheSerLysL 39
                     |||::::::::::   :::   |||       |||:::
     199 CAATGGTGGTGGCGATGAAGGCGACGCTCTGGGCAGTAACAGTTCTTCGC 248

      39 ysLysLysAsnValThrMetArgSerIleLysThrThrArgAspArgVal 55
         ::   ::::::|||:::          :::                ||||||
     249 AACCCAACCGTGTCGCTAGGATGCCAGTT............GATAGAAAT 286

      56 ProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGlyLysCysIleIl 72
            |||   ||||||||||||   :::       |||   :::||||||
     287 GCCCCTTAT...TACAACATGAACCATAAACATCGTGGTATGGCCATTAT 333

      72 eIleAsnAsnLysAsnPheAspLysValThrGlyMetGlyValArgAsnG 89
         |:::|||::::::::::||||||   :::   ::::::      |||:::|
     334 TTTCAACCACGAGCATTTCGAC...ATTCACAGCCTGAAGTCCCGTACTG 380

      89 lyThrAspLysAspAlaGluAlaLeuPheLysCysPheArgSerLeuGly 105
         |||||:::   |||::::::   |||   |||   ::::::::::||||||
     381 GCACGAATGTAGACAGCGACAACCTTTCCAAAGTTCTTAAAACCTTGGGA 430

     106 PheAspValIleValTyrAsnAspCysSerCysAlaLysMetGlnAspLe 122
         |||   |||   |||:::   :::   :::       :::::::::   ::
     431 TTCAAAGTTACGGTGTTCCCTAACTTAAAATCAGAAGAAATCAATAAATT 480

     122 uLeuLysLysAlaSerGluGluAspHisThrAsnAlaAlaCysPheAlaC 139
         ::::::::::::::|||   ||||||::::::|||   |||:::
     481 TATCCAGCAGACTGCAGAGATGGACCATTCAGACGCTGACTGTTTACTTG 530

     139 ysIleLeuLeuSerHisGlyGluGluAsnValIleTyrGlyLysAspGly 155
            :::|||:::||||||||   :::::::::|||:::||||||
     531 TTGCCGTTTTAACCCATGGAGAGCTAGGAATGCTGTATGCCAAAGATACT 580

     156 ValThrProIleLysAspLeuThrAlaHisPheArgGlyAspArgCysLy 172
                  :::|||    :::|||   :::||||:::|||
     581 CATTACAAGCCAGACAACCTGTGGTACTACTTTACTGCTGACAAATGCCC 630

     172 sThrLeuLeuGluLysProLysLeuPhePheIleGlnAlaCysArgGlyT 189
         ||||||      ||||||||||||||||||||||||||||||:::|||
     631 CACACTGGCTGGAAAACCTAAGCTCTTCTTTATTCAGGCTTGCCAAGGTG 680

     189 hrGluLeuAspAspAlaIleGlnAlaAspSerGlyProIleAsnAspThr 205
           :::||||||   :::|||      :::             ::::::|||
     681 ACAGATTGGATGGTGGTATTACTCTGAGCCGC.........ACTGAGACA 721

     206 AspAlaAsnPro......ArgTyrLysIleProValGluAlaAspPheLe 220
         |||::::::|||        |||:::|||||||||:::||||||||||
     722 GATGGCTCACCTTCTACTTCGTATAGGATTCCCGTACATGCTGATTTCTT 771

     220 uPheAlaTyrSerThrValProGlyTyrTyrSerTrpArgSerProGlyA 237
         |:::|||:::|||||||||||||||||||:::||||||||::        |
     772 GATTGCATTTTCAACTGTACCTGGATACTTTTCTTGGAGGAACACTACCA 821

     237 rgGlySerTrpPheValGlnAlaLeuCysSerIleLeuGluGluHisGly 253
         ||||||||||||||:::|||||||||||||:::   |||:::      |||
     822 GAGGTTCATGGTTTATGCAAGCTTTGTGTGAAGAGCTACGCTATGCAGGA 871

     254 LysGluLeuGluIleMetGlnIleLeuThrArgValAsnAspArgValAl 270
            |||   :::|||:::   :::||||||   |||   ::::::||||
     872 ACAGAGAGGGACATTCTGACATTGCTGACTTTTGTGTGTCAGAAGGTTGC 921

     270 aArgHisPheGluSerGlnSer...AspAspProHisPheHisGluLysL 286
         |      ||||||||:::::   |||:::      :::|||:::::::|
     922 GCTAGACTTTGAGTCTAATGCCCCTGACTCAGCGATGATGCATCAACAGA 971

     286 ysGlnIleProCysValValSerMetLeuThrLysGluLeuTyrPheSer 302
         |||||:::||||||:::::|||||||||||||::   |||   |||:::
     972 AACAAGTTCCCTGCATCACCAGCATGCTCACACGTCTGCTTGTGTTTGGT 1021

     303 Gln 303
         :::
    1022 AAG 1024

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-522-813-1

seq_documentation_block:
; Sequence 1, Application US/08522813
; Patent No. 6033848
;  GENERAL INFORMATION:
;    APPLICANT:  Braxton, Scott M
;    APPLICANT:  Diep, Dinh
;    APPLICANT:  Seilhamer, Jeffrey J.
;    TITLE OF INVENTION:  HUMAN ICE HOMOLOG
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3330 Hillview Avenue
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/522,813
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Luther, Barbara J.
;      REGISTRATION NUMBER:  33954
;      REFERENCE/DOCKET NUMBER:  PF-0045P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-852-0195
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1183 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      LIBRARY:  Lung
;      CLONE:  124985
US-08-522-813-1
```

```
alignment_scores:
            Quality:  468.00              Length:    256
              Ratio:  2.644                 Gaps:      1
 Percent Similarity:  69.141   Percent Identity: 37.891

alignment_block:
 US-09-613-508-2 x US-08-522-813-1   ..

 Align seg 1/1  to: US-08-522-813-1  from: 1  to: 1183

      56 ProThrTyrGlnTyrAsnMetAsnPheGluLysLeuGlyLysCysIleIl 72
         |||:::   :::|||:::|||:::  ::::::   |||   ::::::||
      52 CCGGCAGAAAAGTACAAAATGGACCACAGGAGGAGAGGAATTGCTTTAAT 101

      72 eIleAsnAsnLysAsnPheAspLysValThrGlyMetGlyValArgAsnG 89
         |:::|||:::::::::|||                :::      |||:::|
     102 CTTCAATCATGAGAGGTTCTTTTGGCACTTAACACTGCCAGAAAGGCGGG 151

      89 lyThrAspLysAspAlaGluAlaLeuPheLysCysPheArgSerLeuGly 105
         |||||      |||  :::  |||  :::  |||  :::|||||
     152 GCACCTGCGCAGATAGAGACAATCTTACCCGCAGGTTTTCAGATCTAGGA 201

     106 PheAspValIleValTyrAsnAspCysSerCysAlaLysMetGlnAspLe 122
         |||:::|||      :::  |||  ::::::  ::::::
     202 TTTGAAGTGAAATGCTTTTATGATCTTAAAGCAGAAGAACTACTGCTCAA 251

     122 uLeuLysLysAlaSerGluGluAspHisThrAsnAlaAlaCysPheAlaC 139
          :::  :::::::|||      :::|||::::::|||  ||||||:::|
     252 AATTCATGAGGTGTCAACTGTTAGCCACGCAGATGCCGATTGCTTTGTGT 301

     139 ysIleLeuLeuSerHisGlyGluGluAsnValIleTyrGlyLysAspGly 155
         |||::::::|||||||||||||||  |||  ||||||:::  |||:::
     302 GTGTCTTCCTGAGCCATGGCGAAGGCAATCACATTTATGCATATGATGCT 351

     156 ValThrProIleLysAspLeuThrAlaHisPheArgGlyAspArgCysLy 172
              |||:::  |||::::::  |||:::|||||||:::|||
     352 AAAATCGAAATTCAGACATTATCTGGCTCGTTCAAAGGAGACAAGTGTCA 401

     172 sThrLeuLeuGluLysProLysLeuPhePheIleGlnAlaCysArgGlyT 189
          :::|||:::  |||||||||:::|||:::|||||||||||||||||||:
     402 CAGCCTGGTTGGAAAACCCAAGATATTTATCATTCAGGCATGTCGGGGAA 451

     189 hrGlu.............................LeuAspAspAlaIle 195
         :::::                           :::|||:::
     452 ACCAGCACGATGTGCCAGTCATTCCTTTGGATGTAGTAGATAATCAGACA 501

     196 GlnAlaAspSerGlyProIleAsnAspThrAspAlaAsnProArgTyrLy 212
         :::   :::   |||:::::::::|||||||      |||
     502 GAGAAGTTGGACACCAACATAACTGAGGTGGATGCAGCCTCCGTTTACAC 551

     212 sIleProValGluAlaAspPheLeuPheAlaTyrSerThrValProGlyT 229
          :::|||||||  ||||||||||||:::::::|||||:::::   ||||
     552 GCTGCCTGTTGGAGCTGACTTCCTCATGTGTTACTCTGTTGCAGAAGGAT 601

     229 yrTyrSerTrpArgSerProGlyArgGlySerTrpPheValGlnAlaLeu 245
         ||||||||  |||:::      :::|||||||||:::::|||  |||
     602 ATTATTCTCACCGGGAAACTGTGAACGGTTCATGGTACATTCAAGATTTG 651

     246 CysSerIleLeuGluGluHisGlyLysGluLeuGluIleMetGlnIleLe 262
         |||:::::|||  ::::::|||:::::|||||||:::  ::::::||
     652 TGTGAGATGTTGGGAAAATATGGCTCCTCCTTAGAGTTCACAGAACTCCT 701

     262 uThrArgValAsnAspArgValAlaArgHisPheGluSerGlnSerAspA 279
         ||||  ||||||  :::|||:::::::::      :::  :::   |
     702 CACACTGGTGAACAGGAAAGTTTCTCAGCGCCGAGTGGACTTTGGCAAAG 751

     279 spProHisPheHisGluLysLysGlnIleProCysValValSerMetLeu 295
         |||||           |||||||||:::||||||:::  |||||||||
     752 ACCCAAGTGCAATTGGAAAGAAGCAGGTTCCCTGTTTGGGCTCAATGCTA 801

     296 ThrLysGluLeuTyrPhe 301
         ||||||:::|||:::|||
```

```
802 ACTAAAAAGCTGCATTTC 819
```